

STIC-Biotech/Ch mLib

From: Schnizer, Holly
Sent: Monday, June 18, 2001 11:02 AM
To: STIC-Biotech/ChemLib
Subject: seq. search request for appl. no. 09/444,281

Please search the interference and commercial databases for the following sequences:

SEQ ID NO: 35
SEQ ID NO:36 and

a peptide with the following sequence: (SEQ ID NO:35)(SEQ ID NO:27)(SEQ ID NO:35)= peptide of 37 amino acids (A)
a peptide with the following sequence: (SEQ ID NO:36)(SEQ ID NO:27)(SEQ ID NO:35)= peptide of 35 amino acids (B)

Thank you.

Holly Schnizer
AU 1653
CM1-10B05
305-3722

9B01
MB

Holly,

Query B (Seq. IDs 36 + 27 + 35) ~~would create~~
is a peptide = 36 residues.
So, another query was created
wherein Seq. ID: 36 + 27 + 36
was searched. This query
is = 35 residues.

Beverly

Point of Contact:
Beverly Shears
Technical Info. Specialist
CM1 12C14 Tel: 306-4994

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 21, 2001, 08:38:29 ; Search time 42.99 seconds
(without alignments)
52.177 Million cell updates/sec

Title: SCHMIZ-444-MODSEQ1.PEP
Perfect score: 241
Sequence: 1 ILKKPWWPWRRRRHEAPEAPEPIILKKPWWPWRRK 37

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq_0601.*
1: /SID58/gcgdata/geneseq/geneseqp/AA1980.DAT.*
2: /SID58/gcgdata/geneseq/geneseqp/AA1981.DAT.*
3: /SID58/gcgdata/geneseq/geneseqp/AA1982.DAT.*
4: /SID58/gcgdata/geneseq/geneseqp/AA1983.DAT.*
5: /SID58/gcgdata/geneseq/geneseqp/AA1984.DAT.*
6: /SID58/gcgdata/geneseq/geneseqp/AA1985.DAT.*
7: /SID58/gcgdata/geneseq/geneseqp/AA1986.DAT.*
8: /SID58/gcgdata/geneseq/geneseqp/AA1987.DAT.*
9: /SID58/gcgdata/geneseq/geneseqp/AA1988.DAT.*
10: /SID58/gcgdata/geneseq/geneseqp/AA1989.DAT.*
11: /SID58/gcgdata/geneseq/geneseqp/AA1990.DAT.*
12: /SID58/gcgdata/geneseq/geneseqp/AA1991.DAT.*
13: /SID58/gcgdata/geneseq/geneseqp/AA1992.DAT.*
14: /SID58/gcgdata/geneseq/geneseqp/AA1993.DAT.*
15: /SID58/gcgdata/geneseq/geneseqp/AA1994.DAT.*
16: /SID58/gcgdata/geneseq/geneseqp/AA1995.DAT.*
17: /SID58/gcgdata/geneseq/geneseqp/AA1996.DAT.*
18: /SID58/gcgdata/geneseq/geneseqp/AA1997.DAT.*
19: /SID58/gcgdata/geneseq/geneseqp/AA1998.DAT.*
20: /SID58/gcgdata/geneseq/geneseqp/AA1999.DAT.*
21: /SID58/gcgdata/geneseq/geneseqp/AA2000.DAT.*
22: /SID58/gcgdata/geneseq/geneseqp/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	141.5	58.7	28	AA191800	Amino acid sequence
2	137.5	57.1	27	AA191800	Indolicidin analog
3	133.5	55.4	63	AA191800	Poly-(Indol (1-13))
4	133.5	55.4	63	AA191800	Indolicidin fusion
5	99.5	41.3	16	AA191800	Antimicrobial cati
6	99.5	41.3	16	AA191800	Indolicidin analog
7	99.5	41.1	21	AA191800	Amino acid sequence
8	98.5	40.7	21	AA191800	Indolicidin analog
9	98.5	40.7	21	AA191800	Amino acid sequence
10	93.5	38.8	20	AA191800	Indolicidin analog
11	93.5	38.8	20	AA191800	Amino acid sequence

12	91	37.8	13	AA12873	Antimicrobial cati
13	91	37.8	13	AA124609	Indolicidin analog
14	91	37.8	13	AA124609	Cationic peptide o
15	91	37.8	13	AA124609	Cationic peptide M
16	91	37.8	13	AA124609	MBI-11 peptide der
17	91	37.8	13	AA124609	Indolicidin analog
18	91	37.8	13	AA124609	Amino acid sequenc
19	91	37.8	13	AA124609	Amino acid sequenc
20	91	37.8	13	AA124609	Amino acid sequenc
21	91	37.8	13	AA124609	Amino acid sequenc
22	91	37.8	13	AA124609	Amino acid sequenc
23	91	37.8	14	AA124609	Indolicidin analog
24	91	37.8	14	AA124609	Amino acid sequenc
25	89	36.9	15	AA124609	Antimicrobial cati
26	87.5	36.3	21	AA124609	Indolicidin analog
27	87.5	36.3	21	AA124609	Cationic peptide o
28	87.5	36.3	21	AA124609	Amino acid sequenc
29	87	36.1	12	AA124609	Indolicidin analog
30	87	36.1	12	AA124609	Antimicrobial cati
31	86	35.7	12	AA124609	Indolicidin analog
32	86	35.7	12	AA124609	Amino acid sequenc
33	86	35.7	12	AA124609	Indolicidin analog
34	86	35.7	13	AA124609	Antimicrobial cati
35	86	35.7	13	AA124609	Indolicidin analog
36	86	35.7	13	AA124609	Indolicidin analog
37	86	35.7	13	AA124609	Amino acid sequenc
38	86	35.7	13	AA124609	Amino acid sequenc
39	86	35.7	14	AA124609	Indolicidin analog
40	86	35.7	14	AA124609	Antimicrobial cati
41	85.5	35.5	16	AA124609	Amino acid sequenc
42	85	35.3	12	AA124609	Indolicidin analog
43	85	35.3	12	AA124609	Amino acid sequenc
44	85	35.3	13	AA124609	Antimicrobial cati
45	85	35.3	13	AA124609	Antimicrobial cati

ALIGNMENTS

RESULT 1	
AA191800	standard; Peptide; 28 AA.
AA191800;	
06-JUN-2000	(first entry)
XX	Amino acid sequence of cationic peptide MBI 11B20CN.
XX	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
XX	leukaemia; polyoxalylene-modified; APO; lymphoma; multiple myeloma;
XX	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
XX	multidrug resistance.
XX	Synthetic.
XX	WO9965506-A2.
XX	23-DEC-1999.
XX	14-JUN-1999; 99WO-CA00552.
XX	12-JUN-1998; 98US-0096541.
XX	(MICR-) MICROLOGIX BIOTECH INC.
XX	Friedland HD, Krieger TJ, Taylor R, Erffle D, Fraser JR, West MHP;
XX	WPI; 2000-223549/19.
XX	Novel pharmaceutical composition containing optionally activated
XX	polyoxalylene-modified cationic peptides, useful for treating tumours

XX Claim 1; Page 15; 94pp; English.
 PS
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SQ Sequence 28 AA;
 QY
 Db Query Match 58.7%; Score 141.5; DB 21; Length 28;
 Best Local Similarity 58.8%; Pred. No. 1.6e-11;
 Matches 20; Conservative 2; Mismatches 1; Indels 11; Gaps 1;
 4 KMPWPMRRRHEAPEPMILKKMPMPWRRK 37
 :||||||| | :|||||||
 3 TWPWPWIRK-----LMTWPWPWIRK 25
 RESULT 2
 AAM66363
 ID AAM66363 standard; peptide; 27 AA.
 XX
 AC AAM66363;
 XX
 DT 12-JAN-1999 (first entry)
 XX
 DE Indolicidin analogue MBI 11B20.
 XX
 KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism;
 KW bacteria; fungus; parasite; virus.
 XX
 OS Bos taurus.
 OS Synthetic.
 XX
 PN WO9840401-A2.
 XX
 PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, McNicol PJ, West MHP;
 XX
 DR WPI; 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Claim 1; Page 91; 105pp; English.
 CC
 CC The present sequence represents an indolicidin analogue. The present
 CC invention describes compositions and methods for treating infection,
 CC especially bacterial infections. The compositions and methods use
 CC cationic peptides in combination with an antibiotic agent which are
 CC then administered to a patient to enhance the activity of the antibiotic
 CC agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)
 CC inherent resistance. The combinations of antibiotics and cationic
 CC peptides can provide synergistic activity against a microorganism that

CC is tolerant, inherently resistant, or has acquired resistance to an
 CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 CC parasites and viruses.
 CC
 SQ Sequence 27 AA;
 QY
 Db Query Match 57.1%; Score 137.5; DB 19; Length 27;
 Best Local Similarity 61.8%; Pred. No. 5e-11;
 Matches 21; Conservative 2; Mismatches 0; Indels 11; Gaps 2;
 4 KMPWPMRRRHEAPEPMILKKMPMPWRRK 37
 :||||||| | :|||||||
 3 TWPWPWIRK-----LMTWPWPWIRK 25
 RESULT 3
 AAY44668
 ID AAY44668 standard; Protein; 63 AA.
 XX
 AC AAY44668;
 XX
 DT 18-APR-2000 (first entry)
 XX
 DE Poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)3 protein.
 XX
 KW Crosslinked indolicidin analog; X-indolicidin; poly-Indol 1-13;
 KW stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;
 KW protozoacide; virucide; anti-HIV; human immunodeficiency virus-1;
 KW HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;
 KW Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;
 KW Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;
 KW hexapeptide spacer.
 XX
 OS Synthetic.
 OS Bos sp.
 XX
 FH Key
 FT Region
 FT Location/Qualifiers
 FT 1..5
 FT /label= Enterokinase_recognition_site
 FT 5..6
 FT /label= Enterokinase_cleavage_site
 FT 6..7
 FT /label= Cyanogen_bromide_cleavage_site
 FT 20..21
 FT /label= Cyanogen_bromide_cleavage_site
 FT 25..26
 FT /label= Cyanogen_bromide_cleavage_site
 FT 39..40
 FT /label= Cyanogen_bromide_cleavage_site
 FT 44..45
 FT /label= Cyanogen_bromide_cleavage_site
 FT 58..59
 FT /label= Cyanogen_bromide_cleavage_site
 FT 20..25
 FT /label= Hexapeptide_spacer
 FT 39..44
 FT /label= Hexapeptide_spacer
 FT 58..63
 FT /label= Hexapeptide_spacer
 XX
 PN WO9965510-A1.
 XX
 PD 23-DEC-1999.
 XX
 PR 20-MAY-1999; 99WO-US11165.
 PR 18-JUN-1998; 98US-0099631.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Selsted ME, Osapay K;
 XX

DR WPI: 2000-147133/13.
 DR N-PSDB: AA249764.
 XX
 PT Crosslinked indolicidin analogs with antimicrobial activity against
 PT bacteria, yeast, fungi, protozoa and viruses
 XX
 PS Example 1C: Fig 1, 53pp: English.
 XX
 CC The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
 CC which is a naturally occurring peptide isolated from bovine neutrophils
 CC and has antimicrobial activity. The crosslinked indolicidin
 CC (X-indolicidin) analogs are stable and have antimicrobial activity
 CC against gram positive and negative bacteria (e.g. *Staphylococcus aureus*,
 CC *Escherichia coli* and *Salmonella typhimurium*), yeasts and fungi (e.g.
 CC *Candida albicans*, *Cryptococcus neoformans*), protozoa (e.g. *Giardia*
 CC species and *Acanthamoeba* species), and viruses (e.g. HIV-1).
 CC They can be used for reducing or inhibiting the growth or survival of
 CC microorganisms in an environment e.g. a food or food product, a
 CC solution, an inanimate object comprising a surface, or a mammal.
 CC The present sequence is a protein comprising three
 CC copies of Indol 1-13 each separated by a hexapeptide spacer sequence.
 CC A recombinant construct encoding this sequence was used for the
 CC expression of Indol-homoserine (Hse) analog. The ability of
 CC Indol-Hse analog to maintain antimicrobial activity provides a means to
 CC produce X-indolicidin analog precursors in sufficient quantities.
 CC
 SQ Sequence 63 AA:
 Query Match 55.4%; Score 133.5; DB 21; Length 63;
 Best Local Similarity 64.7%; Pred. No. 4,1e-10;
 Matches 22; Conservative 0; Mismatches 5; Indels 7; Gaps 2;
 QY 4 KPMWPMRRKHEAPEPEIMILK-KPMWPMRR 36
 DB 11 KPMWPMRRM-----ARIAMILPMKPMWPMRR 38
 RESULT 4
 ID AAY57142 standard; Protein: 63 AA.
 AC AAY57142;
 XX
 DT 28-FEB-2000 (first entry)
 XX
 DE Indolicidin fusion peptide amino acid sequence.
 XX
 KW Indolicidin analogue: antimicrobial activity; helminth: bacteria; virus;
 KW treatment: inhibit growth; micro-organism: contact lens solution;
 KW transgenic plant; surgical instrument; yeast; fungi; protozoa.
 OS Synthetic.
 XX
 PN WO9558141-A1.
 PM
 XX 18-NOV-1999.
 PD
 XX 05-MAY-1999; 99WO-US09942.
 PF
 XX 12-MAY-1998; 98US-0076227.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX SeIsted ME;
 PI
 XX WPI: 2000-053028/04.
 DR N-PSDB: AA245123.
 DR
 XX New indolicidin analogues, active against bacteria, yeast, fungi,
 PT protozoa and virus, used for, e.g. treating infections -
 XX
 PS Disclosure: Fig 6; 62pp: English.

XX This is the amino acid sequence of an example of a fusion protein which
 CC consists of an indolicidin analogue linked to another peptide.
 CC Peptides AAY57109-Y57138 and AAY57143-Y57144 are new indolicidin
 CC analogues, which have a homoserine residue and/or a truncated amino
 CC terminal region. The analogues have the following amino acid sequence:
 CC Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-Xaa6-Pro-Xaa6-Xaa6-Pro-Xaa6-Xaa7-Xaa7-Xaa8
 CC where:
 CC Xaa1 = Ile, Leu, Val, Ala, Gly or absent;
 CC Xaa2 = Ile, Leu, Val, Ala, Gly or absent;
 CC Xaa3 = Pro or absent;
 CC Xaa4 = Trp, Phe or absent;
 CC Xaa5 = Arg, Lys or absent;
 CC Xaa6 = Trp or Phe;
 CC Xaa7 = Arg, Lys or absent;
 CC Xaa8 = homoserine (Hse), Met, Met-Xaa9-Met or absent, and
 CC Xaa9 = at least one amino acid;
 CC and further provided that: if Xaa2 is present, Xaa8 = Hse, Met or Met-Xaa9-Met;
 CC and provided that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is
 CC absent, Xaa1 and Xaa2 are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3
 CC are absent; and if Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent.
 CC The indolicidin analogues can be used to create a fusion polypeptide
 CC consisting of the analogue linked to a peptide. The indolicidin
 CC analogues have antimicrobial activity against gram positive bacteria,
 CC gram negative bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1).
 CC They are also active against helminths. The analogues can be used for
 CC reducing or inhibiting growth or survival of a microorganism. They can be
 CC used for treating infections. They can also be included in a liquid such
 CC as water or an aqueous solution, e.g. contact lens solution. The
 CC analogues have potential uses in food products, and in objects such as
 CC the surface of an instrument used to prepare food or to perform surgery.
 CC Transgenic plants or animals useful in the food industry can be produced
 CC by introducing a nucleic acid molecule encoding an indolicidin analogue
 CC into the germline cells of such organisms.
 CC
 SQ Sequence 63 AA:
 Query Match 55.4%; Score 133.5; DB 21; Length 63;
 Best Local Similarity 64.7%; Pred. No. 4,1e-10;
 Matches 22; Conservative 0; Mismatches 5; Indels 7; Gaps 2;
 QY 4 KPMWPMRRKHEAPEPEIMILK-KPMWPMRR 36
 DB 11 KPMWPMRRM-----ARIAMILPMKPMWPMRR 38
 RESULT 5
 ID AAW12899 standard; peptide: 16 AA.
 AC AAW12899;
 XX
 DT 10-DEC-1997 (first entry)
 XX
 DE Antimicrobial cationic peptide Cp-26.
 XX
 KW Bacterial: viral; antitumor; food: preservative; inhibitor; growth;
 KW bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;
 KW antiviral; *Candida albicans*; steriliant; *Salmonella*; *Yersinia*;
 KW *Shigella*.
 OS Synthetic.
 XX
 PN WO9708199-A2.
 PM
 XX 06-MAR-1997.
 PD
 XX 23-AUG-1996; 96WO-IB00996.
 PF
 XX 23-AUG-1995; 95US-0002687.
 PR
 XX (UYBR-) UNIV BRITISH COLUMBIA.
 PA

CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

SO Sequence 21 AA;

Query Match 40.7%; Score 98; DB 21; Length 21;
 Best Local Similarity 64.3%; Pred. No. 3.5e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILKKPMPMPRRKHEAPEAPIMILKK 28
 Db 1 ILKKPMPMPWR-----IMILKK 18

RESULT 10

AAV24570 standard; peptide; 20 AA.

XX AAV24570;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #22.

XX Indolicidin; bacterial infection; photo-oxidised solubiliser;

KM antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;

KW additive; shampoo; soap; insecticide; herbicide; preservative;

XX food; technical material.

OS Synthetic.

PN WO9807745-A2.

PD 26-FEB-1998.

PF 21-AUG-1997; 97MO-US14779.

PR 13-JAN-1997; 97US-0034949.

PR 21-AUG-1996; 96US-0024754.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;

DR WPI; 1998-169090/15.

XX New Indolicidin analogues with antimicrobial activity and related

PT nucleic acid - vectors, transformed cells and antibodies, also

PT conjugates with polyoxalkylene glycol and fatty acid to reduce

PT toxicity, useful therapeutically, as disinfectants etc.

XX Claim 12; Page 89; 129pp; English.

XX AAV24549 to AAV24615 represent indolicidin analogues of formulae

CC (I)-(VIII) containing up to 25 amino acids (aa): R₁X₁X₂X₃B (I), B₁X₁X₂X₃B

CC (II), B₁B₂X₁X₂X₃B (III), B₁X₁X₂X₃B₁B₂B₃ (IV), B₁X₁X₂X₃B₁B₂B₃ (V),

CC (VI), B₁B₂X₁X₂X₃B₁B₂B₃ (VII), L₁X₁X₂X₃B₁B₂B₃ (VIII), L₁X₁X₂X₃B₁B₂B₃ (IX),

CC (X) where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,

CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;

CC in (VIII) at least 2 X = F or Y. The analogues are used to treat

CC infections caused by bacteria (Gram positive or negative, or anaerobic);

CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or

CC trematodes) or viruses. Typical of very many pathogens that can be

CC controlled are leishmania, trypanosoma, Ascaris lumbricoides, Fasciola

CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus

CC aureus, listeria, Clostridium, rotavirus and papilloma virus. Compounds

CC derived from the analogues may be used similarly; the compounds may

CC also be prepared from antibiotics or antirhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.

SO Sequence 20 AA;

Query Match 38.8%; Score 93.5; DB 19; Length 20;
 Best Local Similarity 60.7%; Pred. No. 1.2e-05;
 Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

OY 1 ILKKPMPMPRRKHEAPEAPIMILKK 28
 Db 1 ILKKPMPMPWR-----IMILKK 17

RESULT 11

AAV91807 standard; peptide; 20 AA.

XX AAV91807;

DT 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11D5CN.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;

KM leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;

KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

XX multidrug resistance.

OS Synthetic.

PN WO9965506-A2.

PD 23-DEC-1999.

PF 14-JUN-1999; 99MO-CA00552.

PR 12-JUN-1998; 98US-0096541.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

DR WPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated

PT polyoxalkylene-modified cationic peptides, useful for treating tumours

PT Disclosure; Page 15; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which

CC can be used in the pharmaceutical composition of the invention. The

CC invention relates to a pharmaceutical composition containing at least one

CC activated polyoxalkylene (APO)-modified cationic peptide. The

CC modification of peptides with APO increases their activity against tumour

CC cells, including those with a multidrug resistant phenotype. The

CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukemia, multiple myeloma, or tumours of breast, lung, ovary,

CC cervix, uterus, skin, prostate, liver and colon.

SO Sequence 20 AA;

Query Match 38.8%; Score 93.5; DB 21; Length 20;
 Best Local Similarity 60.7%; Pred. No. 1.2e-05;

Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;
 QY 1 ILKKPMPWRRKHAEPEAPIMLKK 28
 Db 1 ILKKPMPWRRKHAEPEAPIMLKK 17

RESULT 12

AAW12873
 ID AAW12873 standard; peptide: 13 AA.

AC AAW12873;

DT 10-DEC-1997 (first entry)

DE Antimicrobial cationic peptide CP-11.

KW Bacterial; viral; antitumour; food; preservative; inhibitor; growth;

KM bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;

KW antiviral; Candida albicans; sterility; Salmonella; Yersinia;

Shigella.

OS Synthetic.

PN MO908199-A2.

PD 06-MAR-1997.

PF 23-AUG-1996; 96MO-IB00996.

PR 23-AUG-1995; 95US-0002687.

PA (UYBR-) UNIV BRITISH COLUMBIA.

PI Falla TJ, Gough M, Hancock RW.

DR WPI; 1997-179179/16.

PT Cationic peptide(s) having anti-microbial activity - used for the

PT inhibition of bacterial and viral growth, as an antitumour agent,

PT and as a food preservative

PS Claim 2; Page 65; 89pp; English.

XX The present sequence represents a specifically claimed novel isolated

CC cationic peptide which has antimicrobial activity. The amino acid

CC sequence of antimicrobial cationic peptides (including the present

CC sequence) is selected from: X1X1ProX2X3X2Pro(X2X2Pro)(X2X3(X5)O;

CC X1X1X3X3X4(X5)ProX2X3X3; X1X1X3(PProTTP)uX3X2X5X2X3X2(X5)O;

CC X1X1X3X3X2Pro(X2X2Pro)(X2X3)m; where m = 1-5; n = 1-2; o = 2-5; r

CC = 0-8; u = 0-1; X1 = Ile, Leu, Val, Phe, Tyr, Trp or Met; X2 = Trp or

CC Phe; X3 = Arg or Lys; X4 = Trp or Lys; and X5 = Phe, Trp, Arg, Lys or

CC Pro. The peptides are preferably amidated or carboxymethylated. The

CC peptides may be used in methods for inhibiting the growth of a bacterium

CC or yeast, or for inhibiting an endotoxaemia or sepsis associated

CC disorder in a subject. The peptides have a broad activity against

CC antibiotic resistant bacteria, combined with activity against the

CC medically important fungus Candida albicans. In addition, the peptides

CC are useful as antitumour agents and/or antiviral agents. The peptides

CC may be used as sterilants or preservatives of materials susceptible to

CC microbial or viral contamination, e.g. in processed foods to inhibit

CC Salmonella, Yersinia and Shigella. The peptides are compact and tend to

CC have a unique polypyrrolone type II extended helix structure that permits

CC them to span the membrane with relatively few amino acids. The peptides

CC possess the ability to work synergistically with antibiotics, and in

CC addition, some of them possess anti-endotoxin activity.

CC N.B. The present sequence represents SEQ ID NO:1 in the claims and

CC examples of the specification, but differs slightly from the SEQ ID NO:1

CC in the sequence listing on page 51 of the specification (see AAW217179).

CC Sequence 13 AA;

Query Match 37.8%; Score 91; DB 18; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILKKPMPWRRKHAEPEAPIMLKK 13
 Db 1 ILKKPMPWRRKHAEPEAPIMLKK 13

RESULT 13

AAW24609
 ID AAW24609 standard; peptide: 13 AA.

AC AAW24609;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #61.

KW Indolicidin; bacterial infection; photo-oxidised solubiliser;

KM antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;

KW additive; shampoo; soap; insecticide; herbicide; preservative;

food; technical material.

OS Synthetic.

PN MO9807745-A2.

PD 26-FEB-1998.

PF 21-AUG-1997; 97WO-US14779.

PR 13-JAN-1997; 97US-0034949.

PA 21-AUG-1996; 96US-0024754.

PI (MICR-) MICROLOGIX BIOTECH INC.

DR Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;

DR WPI; 1998-169090/15.

PT New indolicidin analogues with antimicrobial activity and related

PT nucleic acid - vectors, transformed cells and antibodies, also

PT conjugates with polyoxalkylene glycol and fatty acid to reduce

PT toxicity, useful therapeutically, as disinfectants etc.

PS Example 1; Page 32; 129pp; English.

XX AAW24549 to AAW24615 represent indolicidin analogues of formulae

CC (1)-(VII) containing up to 25 amino acids (aa): RX2X2XB (I), BX2X2XB

CC (II), BBX2X2XB (III), BX2X2XBBA(AA)nmLBBAGS (IV), BX2X2XB(AA)nm

CC (V), LBBX2X2XB(X)XNR (VI), LKX2X2XBRRR (VII) and BBX2X2XBBA (VIII).

CC where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,

CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;

CC in (VIII) at least 2 X = F or Y. The analogues are used to treat

CC infections caused by bacteria (gram positive or negative, or anaerobic);

CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or

CC trematodes) or viruses. Typical of very many pathogens that can be

CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola

CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus

CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds

CC derived from the analogues may be used similarly; the compounds may

CC also be prepared from antibiotics or antiarrhythmic agents. The analogues

CC may be used therapeutically or to coat medical devices; also they are

CC useful as surface disinfectants, as additives to shampoo or soaps, as

CC insecticides or herbicides, or as preservatives for foods and technical

CC materials. The analogues are administered by injection, lavage, orally

CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader

CC spectrum of activity than indolicidin and modification as compounds

CC reduces their toxicity.

CC Sequence 13 AA;

Query Match 37.8%; Score 91; DB 19; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRRK 13
 |||||
 DB 1 ILKKPMPMPRRK 13

RESULT 14
 AAW6378
 ID AAW6378 standard; peptide: 13 AA.
 AC AAW6378;
 XX

DT 12-JAN-1999 (first entry)
 XX

DE Cationic peptide of claim 15 #5.
 XX

KM Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KM bacterial infection; tolerance; antibacterial; microorganism;
 KM bacteria; fungus; parasite; virus.
 OS Synthetic.
 XX

PN WO9840401-A2.
 XX

PD 17-SEP-1998.
 XX

PF 10-MAR-1998; 98WO-CA00190.
 XX

PR 25-FEB-1998; 98US-0030619.
 XX

PR 10-MAR-1997; 97US-0040649.
 XX

PR 20-AUG-1997; 97US-0915314.
 XX

PR 26-SEP-1997; 97US-0060099.
 XX

PA (MICR-) MICROLOGIX BIOTECH INC.
 XX

PI Fraser JR, McNICOL PJ, West MHP;
 XX

DR WPI; 1998-520800/44.
 XX

XX New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX

PS Claim 15; Page 93; 105pp; English.
 XX

CC The present sequence represents a specifically claimed cationic peptide
 CC from the present invention. The present invention describes compositions
 CC and methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 CC
 XX Sequence 13 AA;
 SQ

Query Match 37.8%; Score 91; DB 19; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRRK 13
 |||||
 DB 1 ILKKPMPMPRRK 13

RESULT 15
 AAW71690
 ID AAW71690 standard; peptide: 13 AA.
 XX
 AC AAW71690;
 XX

DT 11-JAN-1999 (first entry)
 XX

DE Cationic peptide MB11 (MW 1879).
 XX

KM MB11; cationic peptide; plasmid PK11; small cryptic plasmid;
 KM replication; RepA; vector; RAMP.
 KM
 XX

OS Synthetic.
 XX

PN WO9841636-A2.
 XX

PD 24-SEP-1998.
 XX

PF 16-MAR-1998; 98WO-CA00214.
 XX

PR 14-MAR-1997; 97US-0040722.
 XX

PA (BUR1/) BURIAN J.
 XX

PA (KAYW/) KAY W W.
 XX

PI Burian J, Kay WW;
 XX

DR WPI; 1998-531571/45.
 XX

PT Increasing plasmid copy number in a cell with the repA gene product
 PT - and an small cryptic plasmid ori sequence, useful for high level
 PT expression of e.g. cytokines, antigens or therapeutic proteins
 XX
 PS Example 13; Page 54; 82pp; English.
 XX

CC MB11 is a small (mol.wt. 1879) cationic peptide. DNA encoding
 CC MB11 has been incorporated into vector pR2h-B1, in which the
 CC replication leader (R21) sequence of RepA (see also AAW71686) is
 CC joined to 2 Hpro peptides (see also AAW71692), to provide a
 CC vector for expression of MB11 in host cells. The invention
 CC provides controlled replication plasmid vectors (RAMP vectors)
 CC comprising a replication origin of a small cryptic plasmid and a
 CC gene encoding RepA. The vectors can reach very high levels of
 CC plasmid replication, but are not lethal to the host cell, and can
 CC be used to direct the high level expression of e.g. cytokines,
 CC antigens and therapeutic proteins.
 CC
 XX Sequence 13 AA;
 SQ

Query Match 37.8%; Score 91; DB 19; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRRK 13
 |||||
 DB 1 ILKKPMPMPRRK 13

Search completed: June 21, 2001, 08:38:29
 Job time: 266 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 21, 2001, 08:38:58 ; Search time 25.06 Seconds

(Without alignments)
29.743 Million cell updates/secTitle: SCHNIZ-444-MODSEQ1.PEP
Perfect score: 241
Sequence: 1 ILKKMPWMPWRKHEAPEAPILKKMPWMPWRK 37Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:
1: /cgn2_6/prodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfilssl.pep:*Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	99.5	41.3	16	4	US-08-702-054B-38
2	99	41.1	21	4	US-08-915-314-54
3	98	40.7	21	4	US-08-915-314-56
4	93.5	38.8	20	4	US-08-915-314-55
5	91	37.8	13	4	US-08-915-314-30
6	91	37.8	13	4	US-08-915-314-62
7	91	37.8	13	4	US-08-915-314-63
8	91	37.8	13	4	US-08-915-314-64
9	91	37.8	13	4	US-08-915-314-57
10	89	36.9	15	4	US-08-702-054B-40
11	87.5	36.3	21	4	US-08-915-314-46
12	87	36.1	12	4	US-08-915-314-52
13	86	35.7	12	4	US-08-915-314-74
14	86	35.7	12	4	US-08-702-054B-5
15	86	35.7	13	4	US-08-915-314-51
16	86	35.7	13	4	US-08-915-314-58
17	86	35.7	13	4	US-08-702-054B-34
18	86	35.7	14	4	US-08-915-314-59
19	85.5	35.5	16	4	US-08-702-054B-11
20	85	35.3	12	4	US-08-915-314-69
21	85	35.3	13	4	US-08-915-314-38
22	85	35.3	13	4	US-08-915-314-45
23	85	35.3	13	4	US-08-702-054B-1
24	85	35.3	13	4	US-08-702-054B-17
25	85	35.3	13	4	US-08-702-054B-32
26	83.5	34.6	15	4	US-08-702-054B-39
27	83	34.4	12	4	US-08-915-314-24

28	83	34.4	13	4	US-08-915-314-49	Sequence 49, Appl
29	83	34.4	13	4	US-08-915-314-50	Sequence 50, Appl
30	83	34.4	13	4	US-08-702-054B-30	Sequence 30, Appl
31	83	34.4	13	4	US-08-702-054B-31	Sequence 31, Appl
32	83	34.4	13	4	US-08-702-054B-35	Sequence 35, Appl
33	83	34.4	20	4	US-08-915-314-47	Sequence 47, Appl
34	82.5	34.2	21	4	US-08-915-314-48	Sequence 48, Appl
35	82	34.0	13	4	US-08-915-314-25	Sequence 25, Appl
36	82	34.0	13	4	US-08-915-314-66	Sequence 66, Appl
37	82	34.0	13	4	US-08-915-314-67	Sequence 67, Appl
38	82	34.0	13	4	US-08-702-054B-33	Sequence 33, Appl
39	81	33.6	11	4	US-08-915-314-75	Sequence 75, Appl
40	80	33.2	14	4	US-08-702-054B-18	Sequence 18, Appl
41	80	33.2	15	4	US-08-702-054B-41	Sequence 41, Appl
42	80	33.2	16	4	US-08-702-054B-2	Sequence 2, Appl
43	79	32.8	17	4	US-08-702-054B-42	Sequence 42, Appl
44	78	32.4	11	4	US-08-915-314-28	Sequence 28, Appl
45	78	32.4	12	4	US-08-915-314-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1
US-08-702-054B-38
Sequence 38, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
US-08-702-054B-38

Query Match 41.3% Score 99.5; DB 4; Length 16;
Best Local Similarity 43.2% Pred. No. 8.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

Db 1 ILKKWPM-----WPMWPMRRK 16

RESULT 2
US-08-915-314-54

Sequence 54, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4500

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-54

Query Match 41.1%; Score 99; DB 4; Length 21;
Best Local Similarity 64.3%; Pred. No. 1.3e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILKKWPMWPMRRKHEAPEAPIMILKK 28
Db 1 ILKKWPMWPMRRK-----MILKK 18

RESULT 3
US-08-915-314-56

Sequence 56, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4500

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 56:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-56

Query Match 40.7%; Score 98; DB 4; Length 21;
Best Local Similarity 64.3%; Pred. No. 1.8e-06;
Matches 18; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILKKWPMWPMRRKHEAPEAPIMILKK 28
Db 1 ILKKWPMWPMRRK-----MILKK 18

RESULT 4
US-08-915-314-55

Sequence 55, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 55:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-55

Query Match 38.8%; Score 93.5; DB 4; Length 20;
Best Local Similarity 60.7%; Pred. No. 6.2e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 11; Gaps 1;

OY 1 ILKKPMPWMPRRKHAEPEAPIMLKK 28
Db 1 ILKKPMPWMPRR-----MILKK 17

RESULT 5
US-08-915-314-30
Sequence 30, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-30

Query Match 37.8%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWMPRRK 13
Db 1 ILKKPMPWMPRRK 13

RESULT 6
US-08-915-314-62
Sequence 62, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="D-Form of Isoleucine"
US-08-915-314-62

Query Match 37.8%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWMPRRK 13
Db 1 ILKKPMPWMPRRK 13

RESULT 7
US-08-915-314-63
Sequence 63, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 622-4900
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note="D-Form of Lysine"
US-08-915-314-63

Query Match 37.8%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWRRK 13
1 ILKKPMPWRRK 13

Db 1 ILKKPMPWRRK 13

RESULT 8
US-08-915-314-64
Sequence 64, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="D-Form of Isoleucine"
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note="D-Form of Lysine"
US-08-915-314-64

Query Match 37.8%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 8.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWRRK 13
1 ILKKPMPWRRK 13

Db 1 ILKKPMPWRRK 13

RESULT 9
US-08-915-314-57
Sequence 57, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-57

Query Match 37.8% Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
|||||
DB 1 ILKKPMPMPRRK 13

RESULT 10
US-08-702-054B-40
Sequence 40, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-054B-40

Query Match 36.9% Score 89; DB 4; Length 15;
Best Local Similarity 43.8%; Pred. No. 1.7e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

QY 1 ILKKPMPMPRRKHEAPEAPIMILKKPMPW 32
|||||
DB 1 ILKKPMPMPRR-----HW 14

RESULT 11
US-08-915-314-46
Sequence 46, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Ertle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

QY 4 KMPMPMPRRKHEAPEAPIMILKK 28
|||||
DB 3 KMPMPMPRRK-----IMILKK 18

RESULT 12
US-08-915-314-52
Sequence 52, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Ertle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Query Match 36.3% Score 87.5; DB 4; Length 21;
Best Local Similarity 60.0%; Pred. No. 3.7e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-52

Query Match 36.1%; Score 87; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LKKPMPMPRRK 13
Db 1 LKKPMPMPRRK 12

RESULT 13
US-08-915-314-74
Sequence 74, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Brille, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-74

Query Match 35.7%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRR 12
Db 1 ILKKPMPMPRR 12

RESULT 14
US-08-702-054B-5
Sequence 5, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-054B-5

Query Match 35.7%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRR 12
Db 1 ILKKPMPMPRR 12

RESULT 15
US-08-915-314-51
Sequence 51, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.

```

1  APPLICANT: Taylor, Robert
2  APPLICANT: Effie, Douglas
3  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
4  TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICININ
5  NUMBER OF SEQUENCES: 90
6  CORRESPONDENCE ADDRESS:
7  ADDRESSEE: SEED AND BERRY LLP
8  STREET: 6300 Columbia Center, 701 Fifth Avenue
9  CITY: Seattle
10 STATE: Washington
11 COUNTRY: USA
12
13 ZIP: 98104
14
15 COMPUTER READABLE FORM:
16 MEDIUM TYPE: Floppy disk
17 COMPUTER: IBM PC compatible
18 OPERATING SYSTEM: PC-DOS/MS-DOS
19 SOFTWARE: PatentIn Release #1.0, Version #1.30
20
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/915,314
23 FILING DATE: 20-AUG-1997
24 CLASSIFICATION: 424
25
26 ATTORNEY/AGENT INFORMATION:
27 NAME: No. 6180604tenburg Ph.D., Carol
28 REGISTRATION NUMBER: 39,317
29 REFERENCE/DOCKET NUMBER: 660081.405
30
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (206) 622-4900
33 TELEFAX: (206) 662-6031
34
35 INFORMATION FOR SEQ ID NO: 51:
36 SEQUENCE CHARACTERISTICS:
37 LENGTH: 13 amino acids
38 TYPE: amino acid
39 STRANDEDNESS:
40 TOPOLOGY: linear
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
96
```

Query Match	35.7%	Score 86:	DB 4:	Length 13:
Best Local Similarity	41.9%	Pred. No.	3.4e-05:	
Matches 13; Conservative	0:	Mismatches	0:	Indels 18; Gaps 1;

```

QY 7 WTPWRRKHEAPEAPETILKKWPWPWRRK 37
    || |||||
Db 1 W-----KKMPWPWRRK 13

```

Search completed: June 21, 2001, 08:38:58
Job time: 215 sec

GenCore version 4.5.
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:39:31 ; Search time 29.25 Seconds
(without alignments)
96.358 Million cell updates/sec

Title: SCHNIZ-444-MODSEQ1.PEP
Perfect score: 241
Sequence: 1 ILKKWPMWPMRKRHEPEEAPRIILKKWPMWPMRKR 37

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	78	32.4	144 1 JC1222	indolicidin precursor
2	63	26.1	192 2 H86543	hypothetical prote
3	63	26.1	192 2 D72081	conserved hypochet
4	62	25.7	314 2 S43916	hypothetical prote
5	59.5	24.7	114 2 T36208	hypothetical prote
6	59.5	24.7	2290 1 GNNYE	genome polyprotein
7	59	24.5	95 2 E86447	protein F5D14.5 (1
8	59	24.5	485 2 S74708	hypothetical prote
9	58	24.1	107 2 T35634	hypothetical prote
10	58	24.1	513 2 S21976	probable RNA-direc
11	57.5	23.9	376 2 T40591	hypothetical prote
12	57	23.7	711 2 C40046	antibiotic transpo
13	56	23.2	452 2 T28094	hypothetical prote
14	55.5	23.0	527 2 S33068	myosin heavy chain
15	55.5	23.0	990 1 G46335	env polyprotein pr
16	55.5	23.0	1940 2 A59287	myosin heavy chain
17	55	22.8	245 2 C69459	transcription regu
18	55	22.8	352 2 S77448	hypothetical prote
19	55	22.8	715 2 B70741	probable moer prot
20	54.5	22.6	134 2 E72532	hypothetical prote
21	54.5	22.6	340 2 PH0217	reverse transcript
22	54.5	22.6	451 2 S30401	hypothetical prote
23	54.5	22.6	547 2 T45635	hypothetical prote
24	54	22.4	144 2 S35331	apidecin 22 precu
25	54	22.4	469 2 B70607	probable cys prot
26	54	22.4	1173 1 VG1HHC	E2 glycoprotein pr
27	53.5	22.2	538 2 B84759	hypothetical prote
28	53.5	22.2	1113 2 JE0315	low-density lipopr
29	53	22.0	68 2 E82799	hypothetical prote

30	53	22.0	406 2 H69143	coenzyme F420-redu
31	53	22.0	424 2 T07742	omega-6 desaturase
32	53	22.0	480 3 JC7552	Shb-like adapter p
33	53	22.0	691 2 D71430	hypothetical prote
34	53	22.0	949 2 E75352	glycine cleavage s
35	53	22.0	2292 2 S35961	capsid polyprotein
36	52.5	21.8	55 2 T11026	H+-transporting AT
37	52.5	21.8	449 2 C84618	hypothetical prote
38	52.5	21.8	471 2 T50016	transcription fact
39	52.5	21.8	490 2 T21365	hypothetical prote
40	52.5	21.8	518 2 T29589	hypothetical prote
41	52.5	21.8	621 2 S37664	peptidomimetic polypro
42	52.5	21.8	630 2 S37663	peptidomimetic polypro
43	52.5	21.8	691 2 G65150	hypothetical 75.1K
44	52.5	21.8	691 2 E86025	hypothetical prote
45	52.5	21.8	962 1 VCLJVS	env polyprotein pr

ALIGNMENTS

RESULT 1

JC1222 indolicidin precursor - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999

C:Accession: JC1222; A42387; S25664

R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.

Biochem. Biophys. Res. Commun. 187, 467-472, 1992

A:Title: CDNA cloning of the neutrophil bactericidal peptide indolicidin.

A:Reference number: JC1222; MUID:92392368

A:Accession: JC1222

A:Molecule type: protein

A:Residues: 131-143 <SEL>

A:Experimental source: neutrophils

A>Note: sequence extracted from NCBI backbone (NCBI:83840)

C:Superfamily: cathelin; cystatin homology

C:Keywords: amidated carboxyl end

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-130/Domain: propeptide #status predicted <PRO>

F:131-143/Product: indolicidin #status experimental <MAT>

F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from follow1

Query Match 32.4% Score 78; DB 1; Length 144;

Best Local Similarity 35.6% Pred. No. 0.0085;

Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 19 EAEPIILKKWPMWPMR 36

126 ELQSVILPKWPMWPMR 143

RESULT 2

H86543 hypothetical protein CP10426 (imported) - Chlamydia pneumoniae (strain J138)

C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae

C>Date: 02-Mar-2001 #sequence-revision 02-Mar-2001 #text-change 02-Mar-2001

C:Accession: H86543

R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T

Nucleic Acids Res. 28, 2311-2314, 2000

A:Title: Comparison of whole genome sequences of Chlamydia pneumoniae J138.

```

A:Reference number: A86491; MUID:20330349
A:Accession: H86543
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <STO>
A:Cross-references: GB:BA000008; NID:g8978798; PIDN:BA98634.1; GSPDB:GN00142
C:Genetics:
A:Gene: CPJ0426

Query Match          26.1%; Score 63; DB 2; Length 192;
Best Local Similarity 42.3%; Pred. No. 0.8;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 5 WPMWPMRRKHAEPEAPEIMILKKMP 30
      |||:::|::|::|::|::|
Db 138 WPMLEPKRKRIEKLPRGEGICFLSAYP 163

RESULT 3
D72081
conserved hypothetical protein frameshifted CP0327 [Imported] - Chlamydomophila pneumoniae
N:Alternate names: hypothetical protein CT277 homolog
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
A:Accession: D72081; G81589
R:Kaltman, S.; Mitchell, B.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: D72081
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <ARN>
A:Cross-references: GB:AE001625; GB:AE001363; NID:g4376695; PIDN:AD18570.1; PID:g4376707
A:Experimental source: strain CML029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
M.; et al.
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255
A:Accession: G81589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <REX>
A:Cross-references: GB:AE002195; GB:AE002161; NID:g7189246; PIDN:AAF38182.1; PID:g7189222
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CPn0426; CP0327

Query Match          26.1%; Score 63; DB 2; Length 192;
Best Local Similarity 42.3%; Pred. No. 0.8;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 5 WPMWPMRRKHAEPEAPEIMILKKMP 30
      |||:::|::|::|::|::|
Db 138 WPMLEPKRKRIEKLPRGEGICFLSAYP 163

RESULT 4
S43916
hypothetical protein 3 - Bacillus stearothermophilus
C:Species: Bacillus stearothermophilus
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 15-Oct-1999
A:Accession: S43916
R:Hong, E.C.; Ferencik, T.
Mol. Gen. Genet. 243, 343-352, 1994
A:Title: Molecular cloning of a maltose transport gene from Bacillus stearothermophilus
A:Reference number: S43914; MUID:94247374
A:Accession: S43916
A:Status: preliminary
A:Molecule type: DNA

```

```

A;Residues: 1-314 <L10>
A;Cross-references: EMBL:L13418; NID:g436964; PIDN:AA71981.1; PID:g436967
A;Note: the authors translated the initiation codon GTG for residue 1 as Val

Query Match      25.7%; Score 62; DB 2; Length 314;
Best Local Similarity 34.7%; Pred. No. 1.8;
Matches 17; Conservative 5; Mismatches 13; Indels 14; Gaps 3;

QY      2 LKKMPMPMRKHEAPEA-----EP-----IMILKKMPWP-----NR 36
          ||| : : : ||| : : | |||
Db       222 LKOWEQAKARWEAREADGAREPGETWDELTKDPSMERLLGRMR 270

RESULT      5
T36208
hypothetical protein SCE36_09 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36208
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21601
A:Accession: T36208
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1114 <OL1>
A:Cross-references: EMBL:A1049763; PIDD:CABA2078.1; GSPDB:GN00070; SCOEDB:SCE36_09
C:Genetics:
A:gene: SCOEDB:SCE36_09

Query Match      24.7%; Score 59.5; DB 2; Length 114;
Best Local Similarity 50.0%; Pred. No. 1.3;
Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY      18 PEAEPIMLIKKW-PMPMPWR 36
          ||| : : : |||||
Db       93 PETAPADARRMRPRMWR 112

RESULT      6
GNMYE
genome polyprotein - encephalomyocarditis virus
N:Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; C
EC 3.4.-.-); RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: encephalomyocarditis virus, EMCV
A;Note: host Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
C:Accession: A03906; JN0383
R:Palmenberg, A.C.; Kirby, E.M.; Janda, M.R.; Drake, N.L.; Duke, G.M.; Potratz, K.F.;
Nucleic Acids Res. 12, 2969-2985, 1984
A>Title: The nucleotide and deduced amino acid sequences of the encephalomyocarditis
A:Reference number: A03906; MUID:84169586
A:Accession: A03906
A:Molecule type: genomic RNA
A:Residues: 1-2290 <PAL>
A:Cross-references: GB:X00463; NID:g61034; PIDD:CAA2152.1; PID:g61035
R:Petrov, N.A.; Chizhikov, V.E.; Bilnov, V.M.; Kartynov, V.A.; Mikryukov, N.N.; Guttor
Blotom. Khim. 10, 274-279, 1984
A>Title: Nucleotide sequence of the 3'-terminus of encephalomyocarditis virus RNA.
A:Reference number: JN0383; MUID:85022788
A:Accession: JN0383
A:Molecule type: genomic RNA
A:Residues: 1337-1396; 'L', 1398-1517, 'A', 1519-1536, 'E', 1538-1556, 'S', 1558-1611, 'T', 1611
A:Cross-references: GB:M54935
A;Note: the authors translated the codon CAU for residue 713 as Thr and AAC for residu
C:Superfamily: foot-and-mouth disease virus genome polyprotein
C:Keywords: coat protein; core protein; genom-linked protein; hydrolase; nucleotidylyl
F:1-67/Domain: leader peptide #status predicted <LP>
F:68-136/Product: coat protein VP4 #status predicted <VP4>
F:137-391/Product: coat protein VP2 #status predicted <VP2>

```

F:392-622/Product: coat protein VP3 #status predicted <VP3>
 F:623-910/Product: coat protein VP1 #status predicted <VP1>
 F:911-1056/Product: core protein P2-A #status predicted <P2A>
 F:1057-1192/Product: core protein P2-B #status predicted <P2B>
 F:1193-1517/Product: core protein P2-C #status predicted <P2C>
 F:1518-1605/Product: core protein P3-A #status predicted <P3A>
 F:1606-1625/Product: genome-linked protein VPg #status predicted <VPg>
 F:1626-1830/Product: proteinase #status predicted <PTS>
 F:1831-2290/Product: RNA-directed RNA polymerase #status predicted <RDP>

Query Match 24.7% Score 59.5; DB 1; Length 2290;
 Best Local Similarity 26.2%; Pred. No. 29;
 Matches 11; Conservative 6; Mismatches 8; Indels 17; Gaps 1;

QY 6 PMWPMRRKHEAPEAEPI-----MILKKP 30
 |||||
 Db 967 PMWPMKNTYQAVLRAPCRVMDIYKKVRFLPLVQKEMP 1008

RESULT 7
 E86447
 Protein F5D14.5 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86447
 R:Releoglas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.;
 anson, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marzalli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: E86447
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-95 <SP0>
 A:Cross-references: GB:AE005172; NID:98920603; PIDN:AAF81325.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F5D14.5
 A:Map position: 1

Query Match 24.5% Score 59; DB 2; Length 95;
 Best Local Similarity 21.3%; Pred. No. 1.2;
 Matches 10; Conservative 5; Mismatches 4; Indels 28; Gaps 3;

OY 5 WP-----WMPMRKHEAPEAEPIIMLK-----KMPMPW 34
 |||||
 Db 46 WPVVVVAVGCGGGRMMW-----PVLVITDVGGEMSWMW 81

RESULT 8
 S74708
 hypothetical protein slr1306 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S74708
 R:Kaneh, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
 O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
 S.
 A:Reference number: S74322; MUID:97061201
 A:Accession: S74708
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-485 <KAN>

A:Cross-references: EMBL:D990901; GB:AB001339; NID:q1651897; PIDN:BA16859.1; PTD:di
 A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 24.5% Score 59; DB 2; Length 485;
 Best Local Similarity 29.3%; Pred. No. 6.6;
 Matches 12; Conservative 6; Mismatches 5; Indels 18; Gaps 2;

OY 6 PW-----WMPMRKHEAPEAEPIIMLK 28
 |||||
 Db 41 PWDGWLWALSGLVIRWRRRHRAPEQOMLPEVLQK 81

RESULT 9
 T35634
 hypothetical protein SC6G9.38 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C>Date: 05-Nov-1989 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999
 C:Accession: T35634
 R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
 submitted to the EMBL Data Library, June 1999
 A:Reference number: 221584
 A:Accession: T35634
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-107 <SDE>
 A:Cross-references: EMBL:AL079356; PIDN:CA845629.1; GSPDB:GN00070; SCOEDB:SC6G9.38
 A:Experimental source: strain A3(2)
 C:Genetics:
 A:Gene: SCOEDB:SC6G9.38

Query Match 24.1% Score 58; DB 2; Length 107;
 Best Local Similarity 50.0%; Pred. No. 1.8;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 3 KKWPMWPMRRKHEAPEAEPI 22
 |||||
 Db 79 RKWPMKRPSPSAEPDTP 98

RESULT 10
 S21976
 Probable RNA-directed DNA polymerase (EC 2.7.7.49) (clone MH2C) - rat retrotranspos
 N:Alternate names: reverse transcriptase
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 19-Mar-1997 #sequence_revision 17-Oct-1997 #text_change 26-Aug-1999
 C:Accession: S21976
 R:Kahre, O.; Ilves, H.; Speck, M.
 submitted to the EMBL Data Library, August 1991
 A:Reference number: S16783
 A:Accession: S21976
 A:Molecule type: mRNA
 A:Residues: 1-513 <KAN>
 A:Cross-references: EMBL:X61295; NID:956521; PIDN:CAA3593.1; PID:956522
 A:Experimental source: clone MH2C
 C:Genetics:
 A:Mobile element: retrotransposon L1
 C:Superfamily: pol polyprotein
 C:Keywords: nucleotidyltransferase; polypeptide; reverse transcriptase

Query Match 24.1% Score 58; DB 2; Length 513;
 Best Local Similarity 39.4%; Pred. No. 9.3;
 Matches 13; Conservative 4; Mismatches 12; Indels 4; Gaps 2;

OY 1 ILKKPMWPMW---RKHEAPEAEPIIMLK-KW 29
 |||||
 Db 373 IFSKWCFMNRATCRBMQIDPCLSPCTKLSKM 405

RESULT 11
 T40591

hypothetical protein SPBC646.15c - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000

C:Accession: T40591

R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: 221938

A:Accession: T40591

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-376 <SER>

A:Cross-references: EMBL:AL035216; PIDN:CAA22819.1; GSPDB:GN00067; SPDB:SPBC646.15c

A:Experimental source: strain 972h; cosmid c646

C:Genetics:

A:Gene: SPDB:SPBC646.15c

A:Map position: 2

A:Introns: 49/1; 126/2; 312/2; 350/1

C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC646.15c

Query Match 23.9%; Score 57.5; DB 2; Length 376;

Best Local Similarity 37.0%; Pred. No. 7; 7; Indels 7; Gaps 1;

Matches 10; Conservative 3; Mismatches 7; Gaps 1;

OY 8 WPMRRKHEAPEEPIIMILKKMPMPW 34

DB 236 WMMRRKXSS-----LKVRRMPW 255

RESULT 12

C40046

antibiotic transport-associated protein actII-3 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Oct-1999

C:Accession: C40046

R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.

Cell 66, 769-780, 1991

A:Title: The act cluster contains regulatory and antibiotic export genes, direct targets

A:Reference number: A40046; MUID:91347376

A:Accession: C40046

A:Molecule type: DNA

A:Residues: 1-711 <FER>

A:Cross-references: GB:M64683; NID:9153143; PIDN:AAA6691.1; PID:9153146

Query Match 23.7%; Score 57; DB 2; Length 711;

Best Local Similarity 47.4%; Pred. No. 17;

Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 ILKKMPMPMRKHEAPE 19

DB 332 IFGRWVFWPAPRKHGTEPD 350

RESULT 13

T28094

hypothetical protein ZK899.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000

C:Accession: T28094

R:Kershaw, J.

submitted to the EMBL Data Library, September 1994

A:Reference number: Z20468

A:Accession: T28094

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-452 <WIL>

A:Cross-references: EMBL:Z37140; PIDN:CAA85502.1; GSPDB:GN00028; CESP:ZK899.2

A:Experimental source: clone ZK899

C:Genetics:

A:Gene: CESP:ZK899.2

A:Map position: X

A:Introns: 34/3; 143/2; 227/2; 262/3; 380/3

C:Superfamily: Caenorhabditis elegans hypothetical protein ZK899.2

Query Match 23.2%; Score 56; DB 2; Length 452;

Best Local Similarity 32.3%; Pred. No. 14;

Matches 10; Conservative 3; Mismatches 10; Indels 8; Gaps 2;

OY 4 KPMWPMRRKHEAPEEPIIMILKKMPMPW 34

DB 172 KLVWMTW---HDTDPN---IFDRMNVW 194

RESULT 14

S33068

myosin heavy chain - fluke (Schistosoma mansoni) (fragment)

N:Alternate names: surface antigen, 200K

C:Species: Schistosoma mansoni

C:Date: 22-Nov-1993 #sequence_revision 06-Sep-1996 #text_change 13-Feb-1998

C:Accession: S33068

R:Solomon, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M.

J. Immunol. 149, 3612-3620, 1992

A:Title: Induction of protective immunity in mice using a 62-kDa recombinant fragment

A:Reference number: A46514; MUID:93056536

A:Accession: S33068

A:Molecule type: mRNA

A:Residues: 1-527 <SOI>

A:Cross-references: EMBL:X65591

A>Note: the authors translated the codon CAA for residue 346 as Lys

C:Superfamily: myosin heavy chain; myosin motor domain homology

C:Keywords: ATP; surface antigen

Query Match

Best Local Similarity 23.0%; Score 55.5; DB 2; Length 527;

Matches 12; Conservative 7; Mismatches 9; Indels 11; Gaps 2;

OY 1 ILKKMPW-----PW---RRKHEAPEEPIIMILKK 28

DB 106 VLRNMPWMLYTKVKPMLNINAROEEMKNAEELARKE 144

RESULT 15

G46335

env polypeptide precursor - Maedi/Visna virus (strain SA-OMV)

N:Alternate names: coat polypeptide

N:Contains: exterior membrane glycoprotein; transmembrane glycoprotein

C:Species: Maedi/Visna virus

C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 16-Jul-1999

C:Accession: G46335

R:Querat, G.; Audoly, G.; Sonigo, P.; Vigne, R.

Virology 175, 434-447, 1990

A:Title: Nucleotide sequence analysis of SA-OMV, a visna-related ovine lentivirus: p

A:Reference number: A46335; MUID:90223989

A:Accession: G46335

A:Molecule type: DNA

A:Residues: 1-990 <QUE>

A:Cross-references: GB:M31646; NID:9808756; PIDN:AAA66817.1; PID:9332551

C:Genetics:

A:Gene: env

C:Superfamily: visna lentivirus type E retrovirus env polypeptide

C:Keywords: capsid protein; coat protein; glycoprotein; polypeptide; transmembrane pr

F:1-101/Domain: signal sequence #status predicted <SIG>

F:102-662/Product: exterior membrane glycoprotein #status predicted <EXT>

F:663-990/Product: transmembrane glycoprotein #status predicted <TM>

F:663-689/Region: hydrophobic

F:842-863/Domain: transmembrane #status predicted <TM>

F:141.162.207.259.299.363.386.402.413.434.438.469.474.480.490.500.514.526.536.542.550

Query Match 23.0%; Score 55.5; DB 1; Length 990;

Best Local Similarity 20.0%; Pred. No. 38;

Matches 14; Conservative 9; Mismatches 8; Indels 39; Gaps 4;

Oy	3	KKWP	----	WP	-----	WRRKHEAPEAE	-----	PIWI	25
		:	:	:	:	:	:	:	:
Db	174	QEMPTNTYHWP	PLMOMENRQW	KKENEXEYTS	SRNNTKED	IDALLAGT	IRGRCPY	PFAL	233
Oy	26	LK	--	KKWP	33				
		:	:	:	:	:	:	:	:
Db	234	LKCE	EMCW	YP	243				

Search completed: June 21, 2001, 08:39:32
Job time: 158 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: June 21, 2001, 08:44:54 ; Search time 17.73 Seconds

(without alignments)
71.486 Million cell updates/sec

Title: SCHNIZ-444-MODSEQ1.PEP
Perfect score: 241
Sequence: 1 ILKKMPMPMPRRKHEAPEAEPIILKKMPMPMPRRK 37

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	32.4	144	1	INDC_BOVIN
2	62	25.7	314	1	YMA3_BACST
3	59.5	24.7	2290	1	POLG_EMCV
4	57	23.7	711	1	MMLA_STRO
5	55.5	23.0	55	1	ATP8_ANAPL
6	55.5	23.0	990	1	ENV_OMVVS
7	55	22.8	715	1	YD55_MYCTU
8	54	22.4	144	1	AP22_APIME
9	54	22.4	469	1	SYCL_MYCTU
10	54	22.4	1173	1	VG12_CVR22
11	53	22.0	424	1	FD6G_SOYBN
12	52.5	21.8	55	1	ATP8_AYTAM
13	52.5	21.8	257	1	E434_ADE40
14	52.5	21.8	691	1	YHJG_ECOLI
15	52.5	21.8	982	1	ENV_VILV
16	52.5	21.8	983	1	ENV_VILV
17	52.5	21.8	991	1	ENV_VILV2
18	52.5	21.8	1154	1	VG12_IBVD2
19	52.5	21.8	1162	1	VG12_IBVB
20	52.5	21.8	1162	1	VG12_IBVB
21	52.5	21.8	1162	1	VG12_IBVB
22	52.5	21.8	1163	1	VG12_IBVB
23	52	21.6	68	1	Y121_BPT4
24	51.5	21.4	162	1	DSBB_NEIMA
25	51.5	21.4	162	1	DSBB_NEIMA
26	51.5	21.4	418	1	PD6C_BRANA
27	51.5	21.4	443	1	PD6C_BRANA
28	51.5	21.4	989	1	ENV_VILV1
29	51	21.2	126	1	YD43_MYCTU
30	51	21.2	126	1	YD43_MYCTU
31	51	21.2	556	1	MEND_ECOLI
32	51	21.2	2292	1	POLG_EMCV
33	51	21.2	2292	1	POLG_EMCV

34	50.5	21.0	887	1	YRAO_HUMAN
35	50	20.7	79	1	YRAO_BACST
36	50	20.7	150	1	NAA4_BPT7
37	50	20.7	157	1	NUQC_SYNY3
38	50	20.7	255	1	LP61_EIME
39	50	20.7	283	1	AP73_APIME
40	50	20.7	413	1	YBHO_ECOLI
41	50	20.7	465	1	TPSN_MOUSE
42	50	20.7	678	1	Y167_YEAST
43	50	20.7	709	1	GLPK_YEAST
44	50	20.7	967	1	MMLA_MYCTU
45	50	20.7	984	1	SM13_MOUSE

ALIGNMENTS

RESULT	ID	INDC_BOVIN	STANDARD	PRT	144 AA
1	INDC_BOVIN	STANDARD	PRT	144 AA	
AC	P33046:	01-OCT-1993 (Rel. 27, Created)			
DT	01-OCT-1993 (Rel. 27, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	INDOLICIDIN PRECURSOR.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OC	Bovidae; Bovinae; Bos.				
OX	NCBI_Taxid=9913;				
RY	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Bone marrow;				
RX	MEDLINE=92392368; PubMed=1520337;				
RA	del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;				
RT	"cDNA cloning of the neutrophil bactericidal peptide indolicidin.";				
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).				
RN	[2]				
RP	SEQUENCE OF 131-143.				
RC	TISSUE=Neutrophils;				
RX	MEDLINE=92165771; PubMed=1537821;				
RA	Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.,				
RT	Cullor J.S.;				
RT	"Indolicidin, a novel bactericidal tridecapeptide amide from				
RT	neutrophils.";				
CC	J. Biol. Chem. 267:4292-4295(1992).				
CC	- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST				
CC	STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.				
CC	- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.				
CC	- PFM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.				
CC	- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: X67340; CAA47755.1; -				
DR	PIR: J01222; J01222.				
DR	PIR: A42387; A42387.				
DR	InterPro: IPR001894; -				
DR	pfam: PF00666; Cathelicidins; 1.				
DR	PROSITE: PS00946; CATHELICIDINS_1; 1.				
DR	PROSITE: PS00946; CATHELICIDINS_2; 1.				
KW	Antibiotic; Amidation; Signal.				
FT	SIGNAL	1	29		POTENTIAL.
FT	PROPEP	30	130		INDOLICIDIN.
FT	PEPTIDE	131	143		PYRROLIDINE-CARBOXYLIC ACID (BY
FT	MOD_RES	30	30		SIMILARITY).

FT DISULFID 85 96 BY SIMILARITY.
 FT DISULFID 107 124 BY SIMILARITY.
 FT MOD_RES 143 143 AMIDATION (G-144 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 144 AA: 16479 MW; E3B1CB855C09911 CRC64;

Query Match 32.4%; Score 78; DB 1; Length 144;
 Best Local Similarity .55.6%; Pred. No. 0.0033;
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 19 EAEPIWIKKMPMPWR 36
 1 : : : : :
 DB 126 ELQSVILPKMPMPWR 143

RESULT 2
 YMA3_BACST STANDARD; PRT; 314 AA.
 AC 045633;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 35.7 KDA PROTEIN IN MALA 3'REGION (ORF3).
 OS Bacillus stearothermophilus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OC NCBI_TaxID-1422;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 7953;
 RX MEDLINE-94247374; PubMed-8190087;
 RA Long E.C., Ferencik T.;
 RT "Molecular cloning of a maltose transport gene from Bacillus
 stearothermophilus and its expression in Escherichia coli K-12.";
 RL Mol. Genet. 243:343-352(1994).
 CC -1- SIMILARITY: BELONGS TO THE UPF0097 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----
 CC DR EMBL; L13418; AAA71981.1; -
 CC KW Hypothetical protein.
 CC SEQUENCE 314 AA; 35735 MW; B54E25FD3F72BFC4 CRC64;

Query Match 25.7%; Score 62; DB 1; Length 314;
 Best Local Similarity 34.7%; Pred. No. 0.71; Length 111
 Matches 17; Conservative 5; Mismatches 13; Indels 14; Gaps 3;

OY 2 LKKVWMPRRKHEAPEA-----EP-----IMLKKMPMP-----WRR 36
 1 : : : : :
 DB 222 LKQWQWQAKARWEARDGARGPGETWDFLLDKPSWMEHLIGWRR 270

RESULT 3
 POLG_EMCV STANDARD; PRT; 2290 AA.
 AC P03304;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN (CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 P2A TO P2C; P3A; GENOME-LINKED PROTEIN VPG; PICORNAIN 3C
 P2A TO P2C; P3A; GENOME-LINKED PROTEIN VPG; PICORNAIN 3C
 DE (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D
 DE (EC 2.7.7.48)1;
 OS Encephalomyocarditis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirus.

OX NCBI_TaxID-12104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-84169586; PubMed-6324136;
 RA Palmberg A.C., Kirby E.M., Janda M.R., Drake N.L., Duke G.M.,
 RA Portratz K.F., Collect M.S.;
 RT "The nucleotide and deduced amino acid sequences of the
 encephalomyocarditis viral polyprotein coding region.";
 RL Nucleic Acids Res. 12:2969-2985(1984).
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib.ch).
 CC -----

CC EMBL; X00463; CA25152.1; -
 CC PIR; A03906; GNMYE.
 CC HSP; P12296; IMEC.
 CC MEROPS; C03.009; -
 CC MEROPS; U29.001; -
 CC Interpro: IPR000605;
 CC Interpro: IPR001205;
 CC Interpro: IPR001676;
 CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
 CC Pfam: PF00910; RNA_helicase; 1.
 CC Pfam: PF00073; rhy; 3.
 CC DR Polyprotein; Coat protein; Core protein; Transferase;
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
 CC KW PROPEP 1
 CC FT CHAIN 1 67
 CC FT CHAIN 1 136
 CC FT CHAIN 137 391
 CC FT CHAIN 392 622
 CC FT CHAIN 623 910
 CC FT CHAIN 911 1056
 CC FT CHAIN 1057 1192
 CC FT CHAIN 1193 1517
 CC FT CHAIN 1518 1605
 CC FT CHAIN 1606 1625
 CC FT CHAIN 1626 1830
 CC FT CHAIN 1831 2290
 CC FT LIPID 68
 CC FT ACT SITE 1784 1784
 CC FT ACT SITE 1802 1802
 CC FT SEQUENCE 2290 AA; 255756 MW; 26BC81BB7CF68CB5 CRC64;

Query Match 24.7%; Score 59.5; DB 1; Length 2290;
 Best Local Similarity 26.2%; Pred. No. 10; Length 111
 Matches 11; Conservative 6; Mismatches 8; Indels 17; Gaps 1;

OY 6 PMPWMPRRKHEAPEA-----IMLKKMP 30
 1 : : : : :
 DB 967 PMPWMPRRKHEAPEA-----IMLKKMP 1008

RESULT 4
 MMLA_STRCO STANDARD; PRT; 711 AA.
 AC Q53902;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN ACTII-3.

```

CC      -1- FINION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC      (CF(0) SUBUNIT) OF THE MITOCHONRIAL ATPASE COMPLEX.
CC      -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC      -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; L22476; AAA72037.1; -
DR      InterPro: IPR001421; -
DR      Pfam: PF00895; ATP-synt_8; 1.
KW      Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT      TRANSEM 8 24 POTENTIAL.
SQ      SEQUENCE 55 AA; 6334 MW; 41284ABAB1525B99 CRC64;

Query Match 23.0%; Score 55.5; DB 1; Length 55;
Best Local Similarity 50.0%; Pred. NO. 0.8;
Matches 9; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

Qy      18 PEAEPIMLKKMPW-MPW 34
Db      37 PSKKPSLTKPTPWMPW 54

RESULT 6
ENV_0MVS STANDARD; PRT; 990 AA.
ID      ENV_0MVS
AC      p16899;
DT      01-AUG-1990 (Rel. 15, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      01-JUN-1994 (Rel. 29, Last annotation update)
DE      ENV POLYPROTEIN PRECURSOR (COAT POLYPROTEIN).
GN      ENV.
OC      Ovine lentivirus (strain SA-OMV).
OC      Viruses; Retrovird viruses; Retroviridae; Lentivirus.
OX      NCBI_Taxid=11664;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=90223989; Pubmed=2158181;
RA      Querat G., Audoly G., Sonigo P., Vigne R.;
RT      "Nucleotide sequence analysis of SA-OMV, a vimsa-related ovine
RT      lentivirus: phylogenetic history of lentiviruses.";
RL      Virology 175:434-447(1990).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M34193; AAA6783.1; -
DR      EMBL; M31646; AAA6817.1; -
DR      PIR; G46335; G46335.
DR      HIV; M34193; ENV5OMVVSACG.
DR      InterPro: IPR000328; -
DR      Pfam: PF00517; GP41; 1.
KW      Glycoprotein; Coat protein; Polyprotein; Transmembrane.
FT      PEPTIDE 1 101 LEADER PEPTIDE.
FT      CHAIN 102 662 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT      CHAIN 663 990 TRANSEMEMPANE GLYCOPROTEIN.
FT      TRANSEM 842 863 POTENTIAL.
FT      CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 207 207 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 363 363 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 413 413 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 500 500 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 526 526 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 560 560 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 567 567 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 703 703 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 771 771 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 778 778 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 794 794 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 990 AA; 114498 MW; 2798185855614F3 CRC64;

Query Match 23.0%; Score 55.5; DB 1; Length 990;
Best Local Similarity 20.0%; Pred. No. 14;
Matches 14; Conservative 9; Mismatches 8; Indels 39; Gaps 4;

OY 3 KKPMP---WP-----NRKHEAPEAE-----PIMI 25
DB 174 QEMPNTYHPIPMQENNRQWKEKEKYTSNNKTKEDIDALLAGKIRGRCVYPAL 233
OY 26 LK-KKPMP 33
DB 234 LKCEMCWYP 243

RESULT 7
ID YD55_MCTU STANDARD; PRT; 715 AA.
AC Q11025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHEICAL 78.2 KDA PROTEIN RV1355C.
GN RV1355C OR MTCY02B10.19C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=968295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Badcock K., Bisham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA Horsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z75555; CA99988.1; -.
DR InterPro: IPR00594; -.
DR Pfam: PF00899; Thif_family; 1.
KW Hypothetical protein.
SQ SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;

Query Match 22.8%; Score 55; DB 1; Length 715;
Best Local Similarity 38.1%; Pred. No. 12;
Matches 8; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

OY 18 PEAPIMI-LKKMPMPMR 36
DB 53 PDPDLEAKRWYPMR 73

RESULT 8
ID AP22_APIME STANDARD; PRT; 144 AA.
AC P35581; P11525; P11526;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE APIDACIN PRECURSOR, TYPE 22.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apoecrita; Aculeata;
OC Apidae; Apidae; Apis.
CX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=93223697; PubMed=8467807;
RA Casteels-Josson K., Capaci T., Casteels P., Tempst P.;
RT "Apidaein multipetide precursor structure: a putative mechanism for
RT amplification of the insect antibacterial response."
RL EMBO J. 12:1569-1578(1993).
RN [2]
RP SEQUENCE (APIDACIN IN/IB).
RC TISSUE=Hemolymph;
RC MEDLINE=90005446; PubMed=2676519;
RA Casteels P., Ampe C., Jacobs F., Vaeck M., Tempst P.;
RT "Apidaein: antibacterial peptides from honeybees."
RL EMBO J. 8:2387-2391(1989).
CC -1- FUNCTION: APIDACIN HAVE BACTERICIDAL ACTIVITY, PREDOMINANTLY
CC AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL
CC PROPAGATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X72576; CA51168.1; -.
DR PIR: S05383; S05383;
DR PIR: S06675; S06675;
DR PIR: S35331; S35331;
DR InterPro: IPR001979; -.
DR Pfam: PF00807; Apidaein; 4.
KW Insect immunity; Antibiotic; Hemolymph; Signal; Multigene family;
KW Cleavage on pair of basic residues; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 35 42 APIDACIN IB.
FT PROPEP 43 60 APIDACIN IB.
FT PROPEP 63 70 APIDACIN IB.
FT PEPTIDE 71 88 APIDACIN IB.
FT PROPEP 91 98

```



```

Query Match      21.8%; Score 52.5; DB 1; Length 257;
Best Local Similarity    43.8%; Pred. No. 8;
Matches          7; Conservative 4; Indels 1; Gaps 1;

Oy      18 PEAEFIMILKKWP-WW 32
        | : ::::| |
Db       25 PSCAFEVVLQEMPMWW 40

```

Query Match	21.88;	Score 52.5;	DB 1;	Length 691;
Best Local Similarity	24.58;	Pred. No. 24;		
Matches 12;	Conservative	3;	Mismatches 13;	Indels 21;
				Gaps 2.

RESULT	15
ENV_VILV	
ID	ENV_VILV
AC	P03379
DT	21-JUL-1986 (Rel. 01, Created)
DT	21-JUL-1986 (Rel. 01, Last sequence update)
DT	15-JUL-1998 (Rel. 38, Last annotation update)
DE	ENV. POLYPROTEIN PRECURSOR (COAT POLYPROTEIN).
GN	ENV.
OS	Visna lentivirus (strain 1514).
OC	Viruses: Retroid viruses: Retroviridae: Lentivirus.
CX	NCBI_TaxId=11742;

RX	MEDLINE=85255938; Pubmed=2410140;
RX	Sonigo P., Alizon M., Staskus K., Klatzmann D., Cole S., Danos O.,
RA	Retzel E., Tjolais P., Haase A., Main-Hobson S.;
RT	"Nucleotide sequence of the vlsna lentivirus: relationship to the
RT	AIDS virus.";
RL	Cell 42:369-382(1985).
RM	[2]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=86062965; Pubmed=2824836;
RA	Braun M.J., Clements J.E., Gonda M.A.;
RT	"The vlsna virus genome: evidence for a hypervariable site in the env
RT	gene and sequence homology among lentivirus envelope proteins.";
RL	J. Virol. 61:4046-4054(1987).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/
CC	or send an email to license@isb-slb.ch).
CC	-----
DR	EMBL: M10608; -: NOT_ANNOTATED_CDS.
DR	EMBL: M51543; -: NOT_ANNOTATED_CDS.
DR	EMBL: A15114; CAA01216.1; -.
DR	PIR: A03977; VCLJVS.
DR	HIV: M10608; ENVSVLVCG.
DR	InterPro: IPR000328; -.
DR	Pfam: PF00517; GP41; 1.
KW	Glycoprotein; Coat protein; Polypotein; Transmembrane.
FT	PEPTIDE
FT	1
FT	LEADER PEPTIDE.
FT	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	TRANSMEMBRANE GLYCOPROTEIN.
FT	POTENTIAL.
FT	TRANSMEM
FT	835 851
FT	CARBOHYD
FT	140 140
FT	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD
FT	161 161
FT	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD
FT	206 206
FT	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD
FT	258 258
FT	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD
FT	298 298
FT	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD
FT	364 364
FT	N-LINKED (GLCNAC. . .) (POTENTIAL)

[illegible]

Score 52.5;	DB 1;	Length 982;
Pred. No. 33;		
Mismatches 1;	Indels 14;	Gaps 5;
EMILLKKMPV 31		
LPILIVIMQV 945		

08:44:55

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:42:25 ; Search time 48.4 Seconds
(without alignments)
101.142 Million cell updates/sec

Title: SCHNIZ-444-MODSEQ1.PEP
Perfect score: 241
Sequence: 1 ILKKPMPWRRKHEAPEPEPIMILKKPMPWRRK 37

Scoring table: BLOSOM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: SP_ARCHAEA:*
 - 2: SP_BACTERIA:*
 - 3: SP_FUNGI:*
 - 4: SP_HUMAN:*
 - 5: SP_INVERTEBRATE:*
 - 6: SP_MAMMAL:*
 - 7: SP_MHC:*
 - 8: SP_ORGANELLE:*
 - 9: SP_PHAGE:*
 - 10: SP_PLANT:*
 - 11: SP_RODENT:*
 - 12: SP_UNCLASSIFIED:*
 - 13: SP_VERTEBRATE:*
 - 14: SP_VIRUS:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70.5	29.3	1245	3	Q9Y7V5
2	67.5	28.0	746	14	Q9JH31
3	67	27.8	723	14	Q9DUC4
4	63	26.1	49	14	Q9DTR80
5	63	26.1	192	2	Q9Z8B7
6	63	26.1	748	14	Q9DTR81
7	62	25.7	367	11	Q63778
8	61	25.3	735	14	Q9DUC9
9	61	25.3	985	14	Q98414
10	60	24.9	147	11	Q61427
11	59.5	24.7	2292	14	Q66765
12	59.5	24.7	95	10	Q9LON0
13	59	24.5	485	10	P72844
14	58	24.1	107	2	Q9XAE4
15	58	24.1	513	11	Q63289
16	58	24.1	646	11	Q63779
17	58	24.1	1300	11	P97692
18	57.5	23.9	117	14	Q9YR89
19					

20	57.5	23.9	341	13	Q90644	Q90644 gallus gall
21	57.5	23.9	376	3	Q94516	Q94516 schistosom
22	57.5	23.9	444	4	Q9HC40	Q9HC40 homo sapien
23	57	23.7	252	14	Q9I037	Q9I037 human immun
24	57	23.7	970	11	Q88821	Q88821 mus musculu
25	57	23.7	971	11	Q70458	Q70458 mus musculu
26	56.5	23.4	157	5	Q9Y0E8	Q9Y0E8 drosophila
27	56.5	23.4	162	5	Q9Y0E8	Q9Y0E8 drosophila
28	56	23.2	252	14	Q9ITX8	Q9ITX8 human immun
29	56	23.2	252	14	Q9ITV2	Q9ITV2 human immun
30	56	23.2	252	14	Q9ITV1	Q9ITV1 human immun
31	56	23.2	252	14	Q9ITV7	Q9ITV7 homo sapien
32	56	23.2	273	4	Q9NRV2	Q9NRV2 caenorhabdi
33	56	23.2	452	5	Q23660	Q23660 schistosoma
34	56	23.2	754	4	Q9JH33	Q9JH33 tt virus. o
35	55.5	23.0	41	8	Q9TENS	Q9TENS anas chloro
36	55.5	23.0	148	5	Q26590	Q26590 schistosoma
37	55.5	23.0	528	5	Q26589	Q26589 schistosoma
38	55.5	23.0	528	5	Q9Y57	Q9Y57 schistosoma
39	55.5	23.0	556	14	Q9WNB7	Q9WNB7 caprine art
40	55.5	23.0	556	14	Q9WNB6	Q9WNB6 caprine art
41	55.5	23.0	556	14	Q9WNB5	Q9WNB5 caprine art
42	55.5	23.0	556	14	Q9WNB4	Q9WNB4 caprine art
43	55.5	23.0	557	14	Q9WNB3	Q9WNB3 caprine art
44	55.5	23.0	557	14	Q9WNB2	Q9WNB2 caprine art
45	55.5	23.0	557	14	Q9WNB2	Q9WNB2 caprine art

ALIGNMENTS

RESULT 1
ID Q9Y7V5 PRELIMINARY: PRT: 1245 AA.
AC Q9Y7V5
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE CONDIOSPORE SURFACE PROTEIN.
CN CMPL
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
OX NCBI_TaxID:5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:ATCC 32173;
RA Pylesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
van Montagu M., Herrera Estrella A., Horwitz B.A.;
RT "Developmental regulation of a gene encoding a multidomain
RT conidiospore surface protein of Trichoderma, cmpl.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ133651; CAB40845.1;
SO SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0CDF8 CRC64;

Query Match 29.3%, Score 70.5; DB 3; Length 1245;
Best Local Similarity 29.0%; Pred. No. 1.2;
Matches 9; Conservative 2; Mismatches 9; Indels 11; Gaps 1;

OY 4 KPPMPWRRKHEAPEPEPIMILKKPMPWRRK 34
DB 1185 RMOQSWPERRGRG-----CWOQMSW 1204
RESULT 2
ID Q9JH31 PRELIMINARY: PRT: 746 AA.
AC Q9JH31
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ORF1.
OS TT virus.

```

OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TJN02;
RA Okamoto H.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TJN02;
RA Ukiha M., Okamoto H., Nishizawa T., Tawara A., Takahashi M.,
  Iizuka H., Miyakawa Y., Mayumi M.;
RT "The entire nucleotide sequences of two distinct TT virus (TTV)
  isolates (TJN01 and TJN02) remotely related to the original TTV
  isolates."
RT Arch. Virol. 0:0-0(2000).
RL EMBL: AB028669; BAA94878.1;
DR SEQUENCE 746 AA; 88561 MW; E0B22953AE764E3E CRC64;
SQ

Query Match
Best Local Similarity 33.3%; Score 67.5; DB 14; Length 746;
Pred. No. 1.7;
Matches 11; Conservative 2; Mismatches 5; Indels 15; Gaps 1;

OY 5 WPMWPMRRKHEAPEEPIMLKKWPMWPMRRK 37
   |||||
Db 3 WGMWPMRRR-----RMPARWRRR 20

RESULT 3
O9DUC4 PRELIMINARY; PRT; 723 AA.
ID O9DUC4
AC O9DUC4
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MF-TTV9;
RA Okamoto H.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MF-TTV9;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
  Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
  phylogenetic relatedness."
RT Virology 277:368-378(2000).
RL EMBL: AB041959; BAB19313.1;
DR SEQUENCE 723 AA; 85393 MW; 232D003098766344 CRC64;
SQ

Query Match
Best Local Similarity 34.4%; Score 67; DB 14; Length 723;
Pred. No. 1.9;
Matches 11; Conservative 1; Mismatches 2; Indels 18; Gaps 1;

OY 6 PMPWPMRRKHEAPEEPIMLKKWPMWPMRRK 37
   |||||
Db 2 PMPWPMRR-----WRRWRRR 15

RESULT 4
O9DTR80 PRELIMINARY; PRT; 49 AA.
ID O9DTR80
AC O9DTR80
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

```

```

DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE ORF1 (FRAGMENT).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TJN9;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
  Sai T., Sugai Y.;
RT "TT virus mRNAs detected in the bone marrow cells from an infected
  individual."
RT Biochem. Biophys. Res. Commun. 279:700-707(2000).
RL EMBL: AB050449; BAB19930.1;
DR NON_TER 49
FT SEQUENCE 49 AA; 7225 MW; 1DA6F8F1AB69AA3 CRC64;
SQ

Query Match
Best Local Similarity 30.3%; Score 63; DB 14; Length 49;
Pred. No. 0.4;
Matches 10; Conservative 3; Mismatches 4; Indels 16; Gaps 2;

OY 5 WPMWPMRRKHEAPEEPIMLKKWPMWPMRRK 37
   |||||
Db 3 WTWV-WQRRRR-----WPMWRRR 19

RESULT 5
O928B7 PRELIMINARY; PRT; 192 AA.
ID O928B7
AC O928B7
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE CP277 SIMILARITY (CPJ0426 OR CP0327).
DN CPJ0426 OR CPJ0426 OR CP0327.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CW1029;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
  Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
RL Nat. Genet. 21:385-389(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-J138;
RA MEDLINE-20330349; PubMed-10871362;
RA Shiral M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
  Shira T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
  from Japan and CW1029 from USA."
RL Nucleic Acids Res. 28:2311-2314(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-AR39;
RA MEDLINE-20150255; PubMed-10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
  White O., Hickey E.K., Peterson J., Umayam L.A., Uetzerback T.,
  Barry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,
  Bowman C., Dodson R., Gwinn M., Nelson W., Deboy R., Kolonay J.,
  McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
  pneumoniae AR39."
RT Nucleic Acids Res. 28:1397-1406(2000).
RL EMBL: AE001625; AAD18570.1;
DR EMBL: AP002546; BAA98634.1;
DR EMBL: AE002194; AAF38182.1;
DR TIGR: CP0327;

```

Query Match 25.78; Score 62; DB 11; Length 36/7

RESULT	9		
ID	098414	PRELIMINARY:	PRT: 985 AA.
AC	098414;		
DT	01-FEB-1997 (TREMBLrel. 02. Created)		
DT	01-FEB-1997 (TREMBLrel. 02. Last sequence update)		
DT	01-MAY-2000 (TREMBLrel. 13. Last annotation update)		
DE	ENVELOPE GLYCOPROTEIN.		
GN	ENV.		
OS	Ovine lentivirus.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11663;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=85/34;		
RA	MEDLINE=55155990; PubMed=7834396;		
RT	Woodward T.M., Carlson J.O., la Concha-Bermejo A., Demattini J.C.,		
RT	"Biological and genetic changes in ovine lentivirus strains following		
RT	passage in isogenic twin lambs."		
RT	J. Acquir. Immune Defic. Syndr. Hum. Retroviro. 8:124-133(1995).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=85/34;		
RA	Carlson J.O., Demattini J.C., Mwaengo D.M.,		
RA	Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.		

Query Match 25.78; Score 62; DB 11; Length 36/7

DR Pfam: PF00517; GP41; 1.
 KM Envelope protein.
 SO SEQUENCE 985 AA; 113794 MW; 3197258EDBD3597 CRC64;

Query Match 25.3%; Score 61; DB 14; Length 985;
 Best Local Similarity 23.4%; Pred. No. 14;
 Matches 18; Conservative 7; Mismatches 8; Indels 44; Gaps 5;

QY 1 ILKK-----WPM-----WPKKHEAPEAE----- 21

DB 163 ILKRYKQDPMTYHWPDMQENMKMKENERYKGTNKTEDIDLLAKIRGRC 222

QY 22 ---PIMILK--KMPMP 33

DB 223 VPPFALLKCTKWCWYR 239

RESULT 10
 ID 061427 PRELIMINARY; PRT; 147 AA.

AC 061427;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE PROCOLLAGEN, TYPE I, ALPHA 1 (ALPHA 1 TYPE I COLLAGEN) (FRAGMENT).
 GN COL1 OR COL1A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C57BL/6; TISSUE=LIVER;
 RX MEDLINE=94344105; PubMed=8065328;
 RA Rhodes K., Rippe R.A., Umezawa A., Nehls M., Brenner D.A., Bretzl M.;
 RT "DNA methylation represses the murine alpha 1(I) collagen promoter by
 an indirect mechanism."
 RL Mol. Cell. Biol. 14:5950-5960(1994).
 DR EMBL; X54876; CAA38657.1; -
 DR MGD; MGI:88467; Col1a1.
 DR InterPro: IPR001007; -
 DR Pfam: PF00093; WWC; 1.
 DR PROSITE: PS01208; WWC; 1.
 DR SMART: SM00214; WWC; 1.
 DR SMART: SM00214; WWC; 1.
 FT NON TER 147
 SO SEQUENCE 147 AA; 16652 MW; 9263BF0A91B4307D CRC64;

Query Match 24.9%; Score 60; DB 11; Length 147;
 Best Local Similarity 34.1%; Pred. No. 2.8;
 Matches 14; Conservative 1; Mismatches 8; Indels 18; Gaps 4;

QY 4 KMPMPRRKHEAPEAPIMILKKV-----PW-----WPMR 35

DB 116 RMP--PSTRW-----PWTAVTSWSSWSPAPRPPMPMR 147

RESULT 11

ID 09X8C2 PRELIMINARY; PRT; 114 AA.

AC 09X8C2;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DE HYPOTHETICAL 13.0 KDA PROTEIN.
 GN SCE36.09.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);
 RA Oliver K., Harris D.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.

QY 18 PEAPRIMILKKV-PWMPMR 36

DB 93 PETAPADAAARRRPPMPMR 112

Query Match 24.7%; Score 59.5; DB 2; Length 114;
 Best Local Similarity 50.0%; Pred. No. 2.5;
 Matches 10; Conservative 2; Mismatches 7; Indels 1; Gaps 1;

QY 18 PEAPRIMILKKV-PWMPMR 36

DB 93 PETAPADAAARRRPPMPMR 112

RESULT 12

ID 066765 PRELIMINARY; PRT; 2292 AA.

AC 066765;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DE POLYPROTEIN, COMPLETE CDS.
 DE Encephalomyocarditis virus.
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirus.
 OX NCBI_TaxID=12104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92148946; PubMed=1310768;
 RA Duke G.M., Hoffman M.A., Palmenberg A.C.;
 RT "Sequence and structural elements that contribute to efficient
 encephalomyocarditis virus RNA translation."
 RL J. Virol. 66:1602-1609(1992).
 DR EMBL; M81861; AAA43037.1; -
 DR HSP; P12296; IMC.
 DR InterPro: IPR000605; -
 DR InterPro: IPR001205; -
 DR InterPro: IPR001676; -

DR Pfam: PF00073; rny; 3.
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.

DR Polypeptide.

KW CHAIN 68 137
 FT CHAIN 138 393
 FT CHAIN 394 624
 FT CHAIN 625 901
 FT CHAIN 902 1044
 FT CHAIN 1045 1194
 FT CHAIN 1195 1519
 FT CHAIN 1520 1627
 FT CHAIN 1628 1832
 FT CHAIN 1833 2292
 SO SEQUENCE 2292 AA; 255457 MW; 01C053788BCEFC94 CRC64;

```
OY      Query Match          24.5%; Score 59; DB 2; Length 485;
OC       Best Local Similarity   29.3%; Pred. No. 12;
OX        Matches    12; Conservative     6; Mismatches     5; Indels    18; Gaps    2;

SYNECHOCOCCUS SP. (strain PCC 6803).
Bacteria; Cyanobacteriales; Synechocystis.
NCBI_TaxID=1148;

[1]
SEQUENCE FROM N.A.
MEDLINE-97061201; PubMed-8905231;
Kaneoka T., Sato S., Kozaki H., Tanaka A., Asamizu E., Nakamura Y.,
Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechococcus sp. strain PCC6803. II. Sequence determination of the
RL entire genome and assignment of potential protein-coding regions.";
RNA DNA Res. 3:109-136(1996);
KW EMBL: D90901; BAA1685;.I; -
DR Hypothetical protein.
SQ SEQUENCE 485 AA; 54270 MW; 9C315C9CEBB03D80 CRC64;

Query Match          24.5%; Score 59; DB 2; Length 485;
Best Local Similarity   29.3%; Pred. No. 12;
Matches    12; Conservative     6; Mismatches     5; Indels    18; Gaps    2;

OY      6 PW-----WPRRRRKEAPEAEPEPIM--ILKK 28
            ||| | | | : : : : 
Db         41 PMDWGMLALSGLCGLVIRRRRHHRHPEDQKOMLEPVLOK 81

RESULT 15
O9XAE4 PRELIMINARY; PRT; 107 AA.
AC O9XAE4: Q9XAE4: (Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
PT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE HYPOTHETICAL 11.6 KDA PROTEIN.
SC SC69.38.
GN Streptomyces coelicolor.
OS Bacteria; Firmicutes; Actinobacteria; Actinomycetiales;
OC Actinomycetales; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC Seeger K.J., Harris D.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RC MEDLINE-97000351; PubMed-884336;
RA Redenbach M., Kieser H.M., Denaparte D., Eichner A., Culm J.,
RA Kinash H., Hopwood D.A.;
RT "A set of ordered cosmid clones and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
DR EMBL: AL079356; CAB45629.1; -.
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 11568 MW; 29F29CF939C03A89 CRC64;
```

Fri Jun 22 08:03:11 2001

schmiz-444-modseq1.psp.rspt

Page 6

Db :||| | ||| |
79 RKWPSKPERPSAEPDTP 98

Search completed: June 21, 2001, 08:42:25
Job time: 276 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:38:29 ; Search time 42.99 Seconds
(without alignments)
50.767 Million cell updates/sec

Title: SCHNIZ-444-MODESEQ2.PEP
Perfect score: 236
Sequence: 1 ILKMPWMPWRKHEAPEPEPIMILKKMPWMPWRK

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues
Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: A-Geneseq.0601.*
 - 2: /SID58/gcgdata/geneseq/geneseq/AA1980.DAT.*
 - 3: /SID58/gcgdata/geneseq/geneseq/AA1981.DAT.*
 - 4: /SID58/gcgdata/geneseq/geneseq/AA1982.DAT.*
 - 5: /SID58/gcgdata/geneseq/geneseq/AA1983.DAT.*
 - 6: /SID58/gcgdata/geneseq/geneseq/AA1984.DAT.*
 - 7: /SID58/gcgdata/geneseq/geneseq/AA1985.DAT.*
 - 8: /SID58/gcgdata/geneseq/geneseq/AA1986.DAT.*
 - 9: /SID58/gcgdata/geneseq/geneseq/AA1987.DAT.*
 - 10: /SID58/gcgdata/geneseq/geneseq/AA1988.DAT.*
 - 11: /SID58/gcgdata/geneseq/geneseq/AA1989.DAT.*
 - 12: /SID58/gcgdata/geneseq/geneseq/AA1990.DAT.*
 - 13: /SID58/gcgdata/geneseq/geneseq/AA1991.DAT.*
 - 14: /SID58/gcgdata/geneseq/geneseq/AA1992.DAT.*
 - 15: /SID58/gcgdata/geneseq/geneseq/AA1993.DAT.*
 - 16: /SID58/gcgdata/geneseq/geneseq/AA1994.DAT.*
 - 17: /SID58/gcgdata/geneseq/geneseq/AA1995.DAT.*
 - 18: /SID58/gcgdata/geneseq/geneseq/AA1996.DAT.*
 - 19: /SID58/gcgdata/geneseq/geneseq/AA1997.DAT.*
 - 20: /SID58/gcgdata/geneseq/geneseq/AA1998.DAT.*
 - 21: /SID58/gcgdata/geneseq/geneseq/AA1999.DAT.*
 - 22: /SID58/gcgdata/geneseq/geneseq/AA2000.DAT.*
 - 23: /SID58/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	152.5	64.6	28	AAV91800	Amino acid sequenc
2	148.5	62.9	27	AAV66363	Indolicidin analog
3	130.5	55.3	63	AAV44668	Poly-(Indol (1-13)
4	130.5	55.3	63	AAV57142	Indolicidin fusion
5	98.5	41.7	21	AAV24552	Indolicidin analog
6	98.5	41.7	21	AAV66376	Cationic peptide o
7	98.5	41.7	21	AAV91796	Amino acid sequenc
8	94.5	40.0	15	AAV66360	Indolicidin analog
9	94.5	40.0	15	AAV91784	Amino acid sequenc
10	94	39.8	20	AAV24553	Indolicidin analog
11	94	39.8	20	AAV91797	Amino acid sequenc

12	93.5	39.6	21	AAV24554	Indolicidin analog
13	93.5	39.6	21	AAV91798	Amino acid sequenc
14	91	38.6	13	AAV12873	Antimicrobial catl
15	91	38.6	13	AAV24569	Indolicidin analog
16	91	38.6	13	AAV66378	Cationic peptide o
17	91	38.6	13	AAV71690	Cationic peptide M
18	91	38.6	13	AAV94495	MBI-11 peptide der
19	91	38.6	13	AAV92795	Indolicidin analog
20	91	38.6	13	AAV91773	Amino acid sequenc
21	91	38.6	13	AAV91774	Amino acid sequenc
22	91	38.6	13	AAV91818	Amino acid sequenc
23	91	38.6	13	AAV91819	Amino acid sequenc
24	91	38.6	13	AAV91820	Indolicidin analog
25	91	38.6	14	AAV24583	Amino acid sequenc
26	91	38.6	14	AAV91804	Indolicidin analog
27	91	38.6	21	AAV24582	Amino acid sequenc
28	91	38.6	21	AAV91806	Indolicidin analog
29	87	36.9	12	AAV24580	Amino acid sequenc
30	86	36.4	12	AAV91804	Indolicidin analog
31	86	36.4	12	AAV12877	Antimicrobial catl
32	86	36.4	12	AAV24550	Indolicidin analog
33	86	36.4	12	AAV24550	Indolicidin analog
34	86	36.4	12	AAV94496	Amino acid sequenc
35	86	36.4	12	AAV91791	Indolicidin analog
36	86	36.4	12	AAV91833	Amino acid sequenc
37	86	36.4	13	AAV12896	Antimicrobial catl
38	86	36.4	13	AAV24613	Indolicidin analog
39	86	36.4	13	AAV24572	Indolicidin analog
40	86	36.4	13	AAV91803	Amino acid sequenc
41	86	36.4	13	AAV91812	Amino acid sequenc
42	86	36.4	14	AAV24573	Indolicidin analog
43	86	36.4	14	AAV91813	Indolicidin analog
44	86	36.4	15	AAV13802	Antimicrobial catl
45	86	36.4	20	AAV24570	Indolicidin analog

ALIGNMENTS

RESULT 1	
AAV91800	AAV91800 standard; Peptide: 28 AA.
ID	
AC	AAV91800;
XX	
XX	06-JUN-2000 (first entry)
DT	
DE	Amino acid sequence of cationic peptide MBI 11B20CN.
XX	
XX	Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
KW	leukaemia; polyoxalkylene-modified; Apo; lymphoma; multiple myeloma;
KW	breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
KW	multidrug resistance.
XX	
OS	Synthetic.
XX	
PN	WO9965506-A2.
XX	
PD	23-DEC-1999.
PF	
XX	14-JUN-1999; 99WO-CA00552.
XX	
PR	12-JUN-1998; 98US-0096541.
XX	
PA	(MICR-) MICROLOGIX BIOTECH INC.
XX	
PI	Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
DR	WPI; 2000-223549/19.
XX	
PT	Novel pharmaceutical composition containing optionally activated
PT	polyoxalkylene-modified cationic peptides, useful for treating tumours

XX Claim 1; Page 15; 94pp; English.

PS

CC This sequence represents a cationic peptide amino acid sequence, which

CC can be used in the pharmaceutical composition of the invention. The

CC invention relates to a pharmaceutical composition containing at least one

CC activated polyoxalyllysine (APO)-modified cationic peptide. The

CC modification of peptides with APO increases their activity against tumour

CC cells, including those with a multidrug resistant phenotype. The

CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,

CC cervix, uterus, skin, prostate, liver and colon.

XX

SO Sequence 28 AA;

Query Match 64.6%; Score 152.5; DB 21; Length 28;

Best Local Similarity 63.9%; Pred. No. 6.3e-13;

Matches 23; Conservative 1; Mismatches 1; Indels 11; Gaps 1;

OY 1 ILRWPMWPMRRKHEAPEPEPMILKKWPMWPMRRK 36

DB 1 ILRWPMWPMRRK-----ILRWPMWPMWPMRRK 25

RESULT 2

AAW6363 standard; peptide; 27 AA.

XX AAW6363;

AC

XX 12-JAN-1999 (first entry)

DT

XX Indolicidin analogue MBI 11B20.

DE

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;

KM bacterial infection; tolerance; antibacterial; microorganism;

KW bacteria; fungus; parasite; virus.

XX

OS Bos taurus.

OS Synthetic.

OS

XX WO9840401-A2.

PN

XX 17-SEP-1998.

PD

XX 10-MAR-1998; 98WO-CA00190.

PF

XX 25-FEB-1998; 98US-0030619.

PR 10-MAR-1997; 97US-0040649.

PR 20-AUG-1997; 97US-0915314.

PR 26-SEP-1997; 97US-0060099.

XX

PA (MICR-) MICROLOGIX BIOTECH INC.

XX

PI Fraser JR, McNicol PJ, West MHP;

XX WPI; 1998-520800/44.

DR

XX New indolicidin peptide analogues - useful for, e.g. enhancing

PT activity of antibiotic or overcoming tolerance, acquired resistance

PT or inherent resistance of microorganisms

XX

PS Claim 1; Page 91; 105pp; English.

XX

CC The present sequence represents an indolicidin analogue. The present

CC invention describes compositions and methods for treating infection,

CC especially bacterial infections. The compositions and methods use

CC cationic peptides in combination with an antibiotic agent which are

CC then administered to a patient to enhance the activity of the antibiotic

CC agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)

CC inherent resistance. The combinations of antibiotics and cationic

CC peptides can provide synergistic activity against a microorganism that

CC is tolerant, inherently resistant, or has acquired resistance to an

CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,

CC parasites and viruses.

XX

SO Sequence 27 AA;

Query Match 62.9%; Score 148.5; DB 19; Length 27;

Best Local Similarity 66.7%; Pred. No. 1.9e-12;

Matches 24; Conservative 1; Mismatches 0; Indels 11; Gaps 2;

OY 1 ILRWPMWPMRRKHEAPEPEPMILKKWPMWPMRRK 36

DB 1 ILRWPMWPMRRK-----MLRWPMWPMWPMRRK 25

RESULT 3

AAW44668 standard; Protein; 63 AA.

XX AAW44668;

AC

XX 18-APR-2000 (first entry)

DT

XX Poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)₃ protein.

DE

XX Crosslinked indolicidin analog; X-indolicidin; poly-indol 1-13;

KM stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;

KW protozoacide; virucide; anti-HIV; human immunodeficiency virus-1;

KW HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;

KW Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;

KW Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;

KW hexapeptide spacer.

XX

OS Synthetic.

OS Bos sp.

OS

XX Key

PH Region

FT Location/Qualifiers

FT 1..5

FT /label- Enterokinase_recognition_site

FT 5..6

FT /label- Enterokinase_cleavage_site

FT 6..7

FT /label- Cyanogen_bromide_cleavage_site

FT 20..21

FT /label- Cyanogen_bromide_cleavage_site

FT 25..26

FT /label- Cyanogen_bromide_cleavage_site

FT 39..40

FT /label- Cyanogen_bromide_cleavage_site

FT 44..45

FT /label- Cyanogen_bromide_cleavage_site

FT 58..59

FT /label- Cyanogen_bromide_cleavage_site

FT 20..25

FT /label- Hexapeptide_spacer

FT 39..44

FT /label- Hexapeptide_spacer

FT 58..63

FT /label- Hexapeptide_spacer

XX

PN WO9965510-A1.

PD

XX 23-DEC-1999.

XX

XX 20-MAY-1999; 99WO-US11165.

PF

XX 18-JUN-1998; 98US-0099631.

PR

XX (REGC) UNIV CALIFORNIA.

PA

XX Selsted ME, Osapay K;

PI

XX

DR WPI: 2000-147133/13.
 DR N-PSDB; AA249764.
 PT Crosslinked indolicidin analogs with antimicrobial activity against
 PT bacteria, yeast, fungi, protozoa and viruses
 XX
 XX Example 1C; Fig 1; 53pp; English.
 CC The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
 CC which is a naturally occurring peptide isolated from bovine neutrophils
 CC and has antimicrobial activity. The crosslinked indolicidin
 CC (X-indolicidin) analogs are stable and have antimicrobial activity
 CC against gram positive and negative bacteria (e.g. *Staphylococcus aureus*,
 CC *Escherichia coli* and *Salmonella typhimurium*), yeasts and fungi (e.g.
 CC *Candida albicans*, *Cryptococcus neoformans*), protozoa (e.g. *Giardia*
 CC *Candida albicans* and *Acanthamoeba* species), and viruses (e.g. HIV-1).
 CC They can be used for reducing or inhibiting the growth or survival of
 CC microorganisms in an environment e.g. a food or food product, a
 CC solution, an inanimate object comprising a surface, or a mammal.
 CC The present sequence is a protein comprising three
 CC copies of Indol 1-13 each separated by a hexapeptide spacer sequence.
 CC A recombinant construct encoding this sequence was used for the
 CC expression of Indol-homoserine (Hse) analog. The ability of
 CC Indol-Hse analog to maintain antimicrobial activity provides a means to
 CC produce X-indolicidin analog precursors in sufficient quantities.
 CC
 SQ Sequence 63 AA:
 3 RWPWPMWRKHEAPEPEPMILK-KWPWMPWR 35
 :||||||| | ||| |||||||
 Db 11 kwpwmpwrirm-----ariamlpwxpwpwr 38
 Query Match 55.3%; Score 130.5; DB 21; Length 63;
 Best Local Similarity 61.8%; Pred. No. 9.2e-10;
 Matches 21; Conservative 1; Mismatches 5; Indels 7; Gaps 2;
 RESULT 4
 ID AAY57142 standard; Protein; 63 AA.
 XX
 AC AAY57142:
 DT 28-FEB-2000 (first entry)
 DE Indolicidin fusion peptide amino acid sequence.
 XX Indolicidin analogue; antimicrobial activity; helminth; bacteria; virus;
 KW treatment; inhibit growth; micro-organism; contact lens solution;
 KW transgenic plant; surgical instrument; yeast; fungi; protozoa.
 XX Synthetic.
 OS
 PN WO9958141-A1.
 XX
 PD 18-NOV-1999.
 XX
 PF 05-MAY-1999; 99WO-US09942.
 XX
 PR 12-MAY-1998; 98US-0076227.
 XX
 XX (REGC) UNIV CALIFORNIA.
 PA
 PI Seleted ME:
 XX
 DR WPI: 2000-053028/04.
 DR N-PSDB; AA245123.
 PT New indolicidin analogues, active against bacteria, yeast, fungi,
 PT protozoa and virus, used for, e.g. treating infections -
 XX
 XX Disclosure; Fig 6; 62pp; English.

XX This is the amino acid sequence of an example of a fusion protein which
 CC consists of an indolicidin analogue linked to another peptide.
 CC Peptides AAY57109-Y57138 and AAY57143-Y57144 are new indolicidin
 CC analogues, which have a homoserine residue and/or a truncated amino
 CC terminal region. The analogues have the following amino acid sequence:
 CC Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-Xaa6-Pro-Xaa6-Pro-Xaa6-Xaa7-Xaa8
 CC where:
 CC Xaa1 = Ile, Leu, Val, Ala, Gly or absent;
 CC Xaa2 = Ile, Leu, Val, Ala, Gly or absent;
 CC Xaa3 = Pro or absent;
 CC Xaa4 = Trp, Phe or absent;
 CC Xaa5 = Arg, Lys or absent;
 CC Xaa6 = Arg or Phe;
 CC Xaa7 = Arg, Lys or absent;
 CC Xaa8 = homoserine (Hse), Met, Met-Xaa9-Met or absent, and
 CC Xaa9 = at least one amino acid;
 CC provided that if Xaa1 is present, Xaa8 = Hse, Met or Met-Xaa9-Met;
 CC and further provided that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is
 CC absent, Xaa1 and Xaa2 are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3
 CC are absent; and if Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent.
 CC The indolicidin analogues can be used to create a fusion polypeptide
 CC consisting of the analogue linked to a peptide. The indolicidin
 CC analogues have antimicrobial activity against gram positive bacteria,
 CC gram negative bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1).
 CC They are also active against helminths. The analogues can be used for
 CC reducing or inhibiting growth or survival of a microorganism. They can be
 CC used for treating infections. They can also be included in a liquid such
 CC as water or an aqueous solution, e.g. contact lens solution. The
 CC analogues have potential uses in food products, and in objects such as
 CC the surface of an instrument used to prepare food or to perform surgery.
 CC Transgenic plants or animals useful in the food industry can be produced
 CC by introducing a nucleic acid molecule encoding an indolicidin analogue
 CC into the germ-line cells of such organisms.
 CC
 SQ Sequence 63 AA:
 3 RWPWPMWRKHEAPEPEPMILK-KWPWMPWR 35
 :||||||| | ||| |||||||
 Db 11 kwpwmpwrirm-----ariamlpwxpwpwr 38
 Query Match 55.3%; Score 130.5; DB 21; Length 63;
 Best Local Similarity 61.8%; Pred. No. 9.2e-10;
 Matches 21; Conservative 1; Mismatches 5; Indels 7; Gaps 2;
 RESULT 5
 ID AAY24552 standard; peptide; 21 AA.
 XX
 AC AAY24552:
 DT 18-AUG-1999 (first entry)
 DE Indolicidin analogue #4.
 XX Indolicidin analogue; bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antiaerhythmic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KW food; technical material.
 XX Synthetic.
 OS
 PN WO9807745-A2.
 XX
 PD 26-FEB-1998.
 XX
 PF 21-AUG-1997; 97WO-US14779.
 XX
 PR 13-JAN-1997; 97US-0034949.
 PR 21-AUG-1996; 96US-0024754.
 XX

PS Disclosure; Page 15; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

SO Sequence 21 AA;

QY 1 ILRPMWPMWRKHEAPEEPIMLKK 27
 |||||
 DB 1 ILRPMWPMWRK-----IMILKK 18

Query Match 41.7%; Score 98.5; DB 21; Length 21;
 Best Local Similarity 66.7%; Pred. No. 3e-06;
 Matches 18; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

RESULT 8
 AAM66360
 ID AAM66360 standard; peptide: 15 AA.
 XX
 AC AAM66360;
 XX
 DT 12-JAN-1999 (first entry)
 DE Indolicidin analogue MBI 11A9.
 XX
 KW Indolicidin analogue; resistance; cationic peptide; antibiotic;
 KW bacterial infection; tolerance; antibacterial; microorganism;
 KW bacteria; fungus; parasite; virus.
 XX
 OS Bos taurus.
 OS Synthetic.
 XX
 PN MO9840401-A2.
 PD 17-SEP-1998.
 XX
 PF 10-MAR-1998; 98WO-CA00190.
 XX
 PR 25-FEB-1998; 98US-0030619.
 PR 10-MAR-1997; 97US-0040649.
 PR 20-AUG-1997; 97US-0915314.
 PR 26-SEP-1997; 97US-0060099.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Fraser JR, McNicol PJ, West MHP;
 XX
 DR WPI: 1998-520800/44.
 XX
 PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Claim 1; Page 91; 105pp; English.
 XX

The present sequence represents an indolicidin analogue. The present
 CC invention describes compositions and methods for treating infection,
 CC especially bacterial infections. The compositions and methods use
 CC cationic peptides in combination with an antibiotic agent which are
 CC then administered to a patient to enhance the activity of the antibiotic
 CC agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)
 CC inherent resistance. The combinations of antibiotics and cationic
 CC peptides can provide synergistic activity against a microorganism that
 CC is tolerant, inherently resistant, or has acquired resistance to an

CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 CC parasites and viruses.

SO Sequence 15 AA;

QY 1 ILRPMWPMWRKHEAPEEPIMLKKWPMWRK 36
 |||||
 DB 1 ILRPMWPMWRK-----WPMWRK 15

Query Match 40.0%; Score 94.5; DB 19; Length 15;
 Best Local Similarity 41.7%; Pred. No. 6.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

RESULT 9
 AAY91784
 ID AAY91784 standard; Peptide: 15 AA.
 XX
 AC AAY91784;
 XX
 DT 06-JUN-2000 (first entry)
 DE Amino acid sequence of cationic peptide MBI 11A9CN.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 XX
 OS Synthetic.
 XX
 PN WO9965506-A2.
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX
 DR WPI: 2000-223549/19.
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours
 XX
 PS Claim 1; Page 14; 94pp; English.
 XX

This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

SO Sequence 15 AA;

QY 1 ILRPMWPMWRKHEAPEEPIMLKKWPMWRK 36
 |||||
 DB 1 ILRPMWPMWRK-----WPMWRK 15

Query Match 40.0%; Score 94.5; DB 21; Length 15;
 Best Local Similarity 41.7%; Pred. No. 6.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

RESULT 10
 AAY24553
 ID AAY24553 standard; peptide; 20 AA.
 AC AAY24553;
 DT 18-AUG-1999 (first entry)
 DE Indolicidin analogue #5.
 XX Indolicidin; bacterial infection; photo-oxidised solubiliser;
 XX antimicrobial; antibacterial; antitumour; surface disinfectant;
 XX additive; shampoo; soap; insecticide; herbicide; preservative;
 XX food; technical material.
 OS Synthetic.
 PN W09807745-A2.
 PD 26-FEB-1998.
 PF 21-AUG-1997; 97WO-US4779.
 PR 13-JAN-1997; 97US-0034949.
 PR 21-AUG-1996; 96US-0024734.
 XX (MICR-) MICROLOGIX BIOTECH INC.
 PA Erle D, Fraser JR, Krieger TJ, Taylor R, West MH;
 PI WPI; 1998-169090/15.
 DR New indolicidin analogues with antimicrobial activity and related
 XX nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 PS Claim 11; Page 88; 129pp; English.
 XX AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (1)-(VIII) containing up to 25 amino acids (aa): R₁X₁X₂X₃X₄ (I), R₁X₁X₂X₃X₄X₅X₆X₇X₈X₉X₁₀X₁₁X₁₂X₁₃X₁₄X₁₅X₁₆X₁₇X₁₈X₁₉X₂₀X₂₁X₂₂X₂₃X₂₄X₂₅ (II), R₁X₁X₂X₃X₄X₅X₆X₇X₈X₉X₁₀X₁₁X₁₂X₁₃X₁₄X₁₅X₁₆X₁₇X₁₈X₁₉X₂₀X₂₁X₂₂X₂₃X₂₄X₂₅X₂₆X₂₇X₂₈X₂₉X₃₀X₃₁X₃₂X₃₃X₃₄X₃₅X₃₆X₃₇X₃₈X₃₉X₄₀X₄₁X₄₂X₄₃X₄₄X₄₅X₄₆X₄₇X₄₈X₄₉X₅₀X₅₁X₅₂X₅₃X₅₄X₅₅X₅₆X₅₇X₅₈X₅₉X₆₀X₆₁X₆₂X₆₃X₆₄X₆₅X₆₆X₆₇X₆₈X₆₉X₇₀X₇₁X₇₂X₇₃X₇₄X₇₅X₇₆X₇₇X₇₈X₇₉X₈₀X₈₁X₈₂X₈₃X₈₄X₈₅X₈₆X₈₇X₈₈X₈₉X₉₀X₉₁X₉₂X₉₃X₉₄X₉₅X₉₆X₉₇X₉₈X₉₉X₁₀₀X₁₀₁X₁₀₂X₁₀₃X₁₀₄X₁₀₅X₁₀₆X₁₀₇X₁₀₈X₁₀₉X₁₁₀X₁₁₁X₁₁₂X₁₁₃X₁₁₄X₁₁₅X₁₁₆X₁₁₇X₁₁₈X₁₁₉X₁₂₀X₁₂₁X₁₂₂X₁₂₃X₁₂₄X₁₂₅X₁₂₆X₁₂₇X₁₂₈X₁₂₉X₁₃₀X₁₃₁X₁₃₂X₁₃₃X₁₃₄X₁₃₅X₁₃₆X₁₃₇X₁₃₈X₁₃₉X₁₄₀X₁₄₁X₁₄₂X₁₄₃X₁₄₄X₁₄₅X₁₄₆X₁₄₇X₁₄₈X₁₄₉X₁₅₀X₁₅₁X₁₅₂X₁₅₃X₁₅₄X₁₅₅X₁₅₆X₁₅₇X₁₅₈X₁₅₉X₁₆₀X₁₆₁X₁₆₂X₁₆₃X₁₆₄X₁₆₅X₁₆₆X₁₆₇X₁₆₈X₁₆₉X₁₇₀X₁₇₁X₁₇₂X₁₇₃X₁₇₄X₁₇₅X₁₇₆X₁₇₇X₁₇₈X₁₇₉X₁₈₀X₁₈₁X₁₈₂X₁₈₃X₁₈₄X₁₈₅X₁₈₆X₁₈₇X₁₈₈X₁₈₉X₁₉₀X₁₉₁X₁₉₂X₁₉₃X₁₉₄X₁₉₅X₁₉₆X₁₉₇X₁₉₈X₁₉₉X₂₀₀X₂₀₁X₂₀₂X₂₀₃X₂₀₄X₂₀₅X₂₀₆X₂₀₇X₂₀₈X₂₀₉X₂₁₀X₂₁₁X₂₁₂X₂₁₃X₂₁₄X₂₁₅X₂₁₆X₂₁₇X₂₁₈X₂₁₉X₂₂₀X₂₂₁X₂₂₂X₂₂₃X₂₂₄X₂₂₅X₂₂₆X₂₂₇X₂₂₈X₂₂₉X₂₃₀X₂₃₁X₂₃₂X₂₃₃X₂₃₄X₂₃₅X₂₃₆X₂₃₇X₂₃₈X₂₃₉X₂₄₀X₂₄₁X₂₄₂X₂₄₃X₂₄₄X₂₄₅X₂₄₆X₂₄₇X₂₄₈X₂₄₉X₂₅₀X₂₅₁X₂₅₂X₂₅₃X₂₅₄X₂₅₅X₂₅₆X₂₅₇X₂₅₈X₂₅₉X₂₆₀X₂₆₁X₂₆₂X₂₆₃X₂₆₄X₂₆₅X₂₆₆X₂₆₇X₂₆₈X₂₆₉X₂₇₀X₂₇₁X₂₇₂X₂₇₃X₂₇₄X₂₇₅X₂₇₆X₂₇₇X₂₇₈X₂₇₉X₂₈₀X₂₈₁X₂₈₂X₂₈₃X₂₈₄X₂₈₅X₂₈₆X₂₈₇X₂₈₈X₂₈₉X₂₉₀X₂₉₁X₂₉₂X₂₉₃X₂₉₄X₂₉₅X₂₉₆X₂₉₇X₂₉₈X₂₉₉X₃₀₀X₃₀₁X₃₀₂X₃₀₃X₃₀₄X₃₀₅X₃₀₆X₃₀₇X₃₀₈X₃₀₉X₃₁₀X₃₁₁X₃₁₂X₃₁₃X₃₁₄X₃₁₅X₃₁₆X₃₁₇X₃₁₈X₃₁₉X₃₂₀X₃₂₁X₃₂₂X₃₂₃X₃₂₄X₃₂₅X₃₂₆X₃₂₇X₃₂₈X₃₂₉X₃₃₀X₃₃₁X₃₃₂X₃₃₃X₃₃₄X₃₃₅X₃₃₆X₃₃₇X₃₃₈X₃₃₉X₃₄₀X₃₄₁X₃₄₂X₃₄₃X₃₄₄X₃₄₅X₃₄₆X₃₄₇X₃₄₈X₃₄₉X₃₅₀X₃₅₁X₃₅₂X₃₅₃X₃₅₄X₃₅₅X₃₅₆X₃₅₇X₃₅₈X₃₅₉X₃₆₀X₃₆₁X₃₆₂X₃₆₃X₃₆₄X₃₆₅X₃₆₆X₃₆₇X₃₆₈X₃₆₉X₃₇₀X₃₇₁X₃₇₂X₃₇₃X₃₇₄X₃₇₅X₃₇₆X₃₇₇X₃₇₈X₃₇₉X₃₈₀X₃₈₁X₃₈₂X₃₈₃X₃₈₄X₃₈₅X₃₈₆X₃₈₇X₃₈₈X₃₈₉X₃₉₀X₃₉₁X₃₉₂X₃₉₃X₃₉₄X₃₉₅X₃₉₆X₃₉₇X₃₉₈X₃₉₉X₄₀₀X₄₀₁X₄₀₂X₄₀₃X₄₀₄X₄₀₅X₄₀₆X₄₀₇X₄₀₈X₄₀₉X₄₁₀X₄₁₁X₄₁₂X₄₁₃X₄₁₄X₄₁₅X₄₁₆X₄₁₇X₄₁₈X₄₁₉X₄₂₀X₄₂₁X₄₂₂X₄₂₃X₄₂₄X₄₂₅X₄₂₆X₄₂₇X₄₂₈X₄₂₉X₄₃₀X₄₃₁X₄₃₂X₄₃₃X₄₃₄X₄₃₅X₄₃₆X₄₃₇X₄₃₈X₄₃₉X₄₄₀X₄₄₁X₄₄₂X₄₄₃X₄₄₄X₄₄₅X₄₄₆X₄₄₇X₄₄₈X₄₄₉X₄₅₀X₄₅₁X₄₅₂X₄₅₃X₄₅₄X₄₅₅X₄₅₆X₄₅₇X₄₅₈X₄₅₉X₄₆₀X₄₆₁X₄₆₂X₄₆₃X₄₆₄X₄₆₅X₄₆₆X₄₆₇X₄₆₈X₄₆₉X₄₇₀X₄₇₁X₄₇₂X₄₇₃X₄₇₄X₄₇₅X₄₇₆X₄₇₇X₄₇₈X₄₇₉X₄₈₀X₄₈₁X₄₈₂X₄₈₃X₄₈₄X₄₈₅X₄₈₆X₄₈₇X₄₈₈X₄₈₉X₄₉₀X₄₉₁X₄₉₂X₄₉₃X₄₉₄X₄₉₅X₄₉₆X₄₉₇X₄₉₈X₄₉₉X₅₀₀X₅₀₁X₅₀₂X₅₀₃X₅₀₄X₅₀₅X₅₀₆X₅₀₇X₅₀₈X₅₀₉X₅₁₀X₅₁₁X₅₁₂X₅₁₃X₅₁₄X₅₁₅X₅₁₆X₅₁₇X₅₁₈X₅₁₉X₅₂₀X₅₂₁X₅₂₂X₅₂₃X₅₂₄X₅₂₅X₅₂₆X₅₂₇X₅₂₈X₅₂₉X₅₃₀X₅₃₁X₅₃₂X₅₃₃X₅₃₄X₅₃₅X₅₃₆X₅₃₇X₅₃₈X₅₃₉X₅₄₀X₅₄₁X₅₄₂X₅₄₃X₅₄₄X₅₄₅X₅₄₆X₅₄₇X₅₄₈X₅₄₉X₅₅₀X₅₅₁X₅₅₂X₅₅₃X₅₅₄X₅₅₅X₅₅₆X₅₅₇X₅₅₈X₅₅₉X₅₆₀X₅₆₁X₅₆₂X₅₆₃X₅₆₄X₅₆₅X₅₆₆X₅₆₇X₅₆₈X₅₆₉X₅₇₀X₅₇₁X₅₇₂X₅₇₃X₅₇₄X₅₇₅X₅₇₆X₅₇₇X₅₇₈X₅₇₉X₅₈₀X₅₈₁X₅₈₂X₅₈₃X₅₈₄X₅₈₅X₅₈₆X₅₈₇X₅₈₈X₅₈₉X₅₉₀X₅₉₁X₅₉₂X₅₉₃X₅₉₄X₅₉₅X₅₉₆X₅₉₇X₅₉₈X₅₉₉X₆₀₀X₆₀₁X₆₀₂X₆₀₃X₆₀₄X₆₀₅X₆₀₆X₆₀₇X₆₀₈X₆₀₉X₆₁₀X₆₁₁X₆₁₂X₆₁₃X₆₁₄X₆₁₅X₆₁₆X₆₁₇X₆₁₈X₆₁₉X₆₂₀X₆₂₁X₆₂₂X₆₂₃X₆₂₄X₆₂₅X₆₂₆X₆₂₇X₆₂₈X₆₂₉X₆₃₀X₆₃₁X₆₃₂X₆₃₃X₆₃₄X₆₃₅X₆₃₆X₆₃₇X₆₃₈X₆₃₉X₆₄₀X₆₄₁X₆₄₂X₆₄₃X₆₄₄X₆₄₅X₆₄₆X₆₄₇X₆₄₈X₆₄₉X₆₅₀X₆₅₁X₆₅₂X₆₅₃X₆₅₄X₆₅₅X₆₅₆X₆₅₇X₆₅₈X₆₅₉X₆₆₀X₆₆₁X₆₆₂X₆₆₃X₆₆₄X₆₆₅X₆₆₆X₆₆₇X₆₆₈X₆₆₉X₆₇₀X₆₇₁X₆₇₂X₆₇₃X₆₇₄X₆₇₅X₆₇₆X₆₇₇X₆₇₈X₆₇₉X₆₈₀X₆₈₁X₆₈₂X₆₈₃X₆₈₄X₆₈₅X₆₈₆X₆₈₇X₆₈₈X₆₈₉X₆₉₀X₆₉₁X₆₉₂X₆₉₃X₆₉₄X₆₉₅X₆₉₆X₆₉₇X₆₉₈X₆₉₉X₇₀₀X₇₀₁X₇₀₂X₇₀₃X₇₀₄X₇₀₅X₇₀₆X₇₀₇X₇₀₈X₇₀₉X₇₁₀X₇₁₁X₇₁₂X₇₁₃X₇₁₄X₇₁₅X₇₁₆X₇₁₇X₇₁₈X₇₁₉X₇₂₀X₇₂₁X₇₂₂X₇₂₃X₇₂₄X₇₂₅X₇₂₆X₇₂₇X₇₂₈X₇₂₉X₇₃₀X₇₃₁X₇₃₂X₇₃₃X₇₃₄X₇₃₅X₇₃₆X₇₃₇X₇₃₈X₇₃₉X₇₄₀X₇₄₁X₇₄₂X₇₄₃X₇₄₄X₇₄₅X₇₄₆X₇₄₇X₇₄₈X₇₄₉X₇₅₀X₇₅₁X₇₅₂X₇₅₃X₇₅₄X₇₅₅X₇₅₆X₇₅₇X₇₅₈X₇₅₉X₇₆₀X₇₆₁X₇₆₂X₇₆₃X₇₆₄X₇₆₅X₇₆₆X₇₆₇X₇₆₈X₇₆₉X₇₇₀X₇₇₁X₇₇₂X₇₇₃X₇₇₄X₇₇₅X₇₇₆X₇₇₇X₇₇₈X₇₇₉X₇₈₀X₇₈₁X₇₈₂X₇₈₃X₇₈₄X₇₈₅X₇₈₆X₇₈₇X₇₈₈X₇₈₉X₇₉₀X₇₉₁X₇₉₂X₇₉₃X₇₉₄X₇₉₅X₇₉₆X₇₉₇X₇₉₈X₇₉₉X₈₀₀X₈₀₁X₈₀₂X₈₀₃X₈₀₄X₈₀₅X₈₀₆X₈₀₇X₈₀₈X₈₀₉X₈₁₀X₈₁₁X₈₁₂X₈₁₃X₈₁₄X₈₁₅X₈₁₆X₈₁₇X₈₁₈X₈₁₉X₈₂₀X₈₂₁X₈₂₂X₈₂₃X₈₂₄X₈₂₅X₈₂₆X₈₂₇X₈₂₈X₈₂₉X₈₃₀X₈₃₁X₈₃₂X₈₃₃X₈₃₄X₈₃₅X₈₃₆X₈₃₇X₈₃₈X₈₃₉X₈₄₀X₈₄₁X₈₄₂X₈₄₃X₈₄₄X₈₄₅X₈₄₆X₈₄₇X₈₄₈X₈₄₉X₈₅₀X₈₅₁X₈₅₂X₈₅₃X₈₅₄X₈₅₅X₈₅₆X₈₅₇X₈₅₈X₈₅₉X₈₆₀X₈₆₁X₈₆₂X₈₆₃X₈₆₄X₈₆₅X₈₆₆X₈₆₇X₈₆₈X₈₆₉X₈₇₀X₈₇₁X₈₇₂X₈₇₃X₈₇₄X₈₇₅X₈₇₆X₈₇₇X₈₇₈X₈₇₉X₈₈₀X₈₈₁X₈₈₂X₈₈₃X₈₈₄X₈₈₅X₈₈₆X₈₈₇X₈₈₈X₈₈₉X₈₉₀X₈₉₁X₈₉₂X₈₉₃X₈₉₄X₈₉₅X₈₉₆X₈₉₇X₈₉₈X₈₉₉X₉₀₀X₉₀₁X₉₀₂X₉₀₃X₉₀₄X₉₀₅X₉₀₆X₉₀₇X₉₀₈X₉₀₉X₉₁₀X₉₁₁X₉₁₂X₉₁₃X₉₁₄X₉₁₅X₉₁₆X₉₁₇X₉₁₈X₉₁₉X₉₂₀X₉₂₁X₉₂₂X₉₂₃X₉₂₄X₉₂₅X₉₂₆X₉₂₇X₉₂₈X₉₂₉X₉₃₀X₉₃₁X₉₃₂X₉₃₃X₉₃₄X₉₃₅X₉₃₆X₉₃₇X₉₃₈X₉₃₉X₉₄₀X₉₄₁X₉₄₂X₉₄₃X₉₄₄X₉₄₅X₉₄₆X₉₄₇X₉₄₈X₉₄₉X₉₅₀X₉₅₁X₉₅₂X₉₅₃X₉₅₄X₉₅₅X₉₅₆X₉₅₇X₉₅₈X₉₅₉X₉₆₀X₉₆₁X₉₆₂X₉₆₃X₉₆₄X₉₆₅X₉₆₆X₉₆₇X₉₆₈X₉₆₉X₉₇₀X₉₇₁X₉₇₂X₉₇₃X₉₇₄X₉₇₅X₉₇₆X₉₇₇X₉₇₈X₉₇₉X₉₈₀X₉₈₁X₉₈₂X₉₈₃X₉₈₄X₉₈₅X₉₈₆X₉₈₇X₉₈₈X₉₈₉X₉₉₀X₉₉₁X₉₉₂X₉₉₃X₉₉₄X₉₉₅X₉₉₆X₉₉₇X₉₉₈X₉₉₉X₁₀₀₀X₁₀₀₁X₁₀₀₂X₁₀₀₃X₁₀₀₄X₁₀₀₅X₁₀₀₆X₁₀₀₇X₁₀₀₈X₁₀₀₉X₁₀₁₀X₁₀₁₁X₁₀₁₂X₁₀₁₃X₁₀₁₄X₁₀₁₅X₁₀₁₆X₁₀₁₇X₁₀₁₈X₁₀₁₉X₁₀₂₀X₁₀₂₁X₁₀₂₂X₁₀₂₃X₁₀₂₄X₁₀₂₅X₁₀₂₆X₁₀₂₇X₁₀₂₈X₁₀₂₉X₁₀₃₀X₁₀₃₁X₁₀₃₂X₁₀₃₃X₁₀₃₄X₁₀₃₅X₁₀₃₆X₁₀₃₇X₁₀₃₈X₁₀₃₉X₁₀₄₀X₁₀₄₁X₁₀₄₂X₁₀₄₃X₁₀₄₄X₁₀₄₅X₁₀₄₆X₁₀₄₇X₁₀₄₈X₁₀₄₉X₁₀₅₀X₁₀₅₁X₁₀₅₂X₁₀₅₃X₁₀₅₄X₁₀₅₅X₁₀₅₆X₁₀₅₇X₁₀₅₈X₁₀₅₉X₁₀₆₀X₁₀₆₁X₁₀₆₂X₁₀₆₃X₁₀₆₄X₁₀₆₅X₁₀₆₆X₁₀₆₇X₁₀₆₈X₁₀₆₉X₁₀₇₀X₁₀₇₁X₁₀₇₂X₁₀₇₃X₁₀₇₄X₁₀₇₅X₁₀₇₆X₁₀₇₇X₁₀₇₈X₁₀₇₉X₁₀₈₀X₁₀₈₁X₁₀₈₂X₁₀₈₃X₁₀₈₄X₁₀₈₅X₁₀₈₆X₁₀₈₇X₁₀₈₈X₁₀₈₉X₁₀₉₀X₁₀₉₁X₁₀₉₂X₁₀₉₃X₁₀₉₄X₁₀₉₅X₁₀₉₆X₁₀₉₇X₁₀₉₈X₁₀₉₉X₁₁₀₀X₁₁₀₁X₁₁₀₂X₁₁₀₃X₁₁₀₄X₁₁₀₅X₁₁₀₆X₁₁₀₇X₁₁₀₈X₁₁₀₉X₁₁₁₀X₁₁₁₁X₁₁₁₂X₁₁₁₃X₁₁₁₄X₁₁₁₅X₁₁₁₆X₁₁₁₇X₁₁₁₈X₁₁₁₉X₁₁₂₀X₁₁₂₁X₁₁₂₂X₁₁₂₃X₁₁₂₄X₁₁₂₅X₁₁₂₆X₁₁₂₇X₁₁₂₈X₁₁₂₉X₁₁₃₀X₁₁₃₁X₁₁₃₂X₁₁₃₃X₁₁₃₄X₁₁₃₅X₁₁₃₆X₁₁₃₇X₁₁₃₈X₁₁₃₉X₁₁₄₀X₁₁₄₁X₁₁₄₂X₁₁₄₃X₁₁₄₄X₁₁₄₅X₁₁₄₆X₁₁₄₇X₁₁₄₈X₁₁₄₉X₁₁₅₀X₁₁₅₁X₁₁₅₂X₁₁₅₃X₁₁₅₄X₁₁₅₅X₁₁₅₆X₁₁₅₇X₁₁₅₈X₁₁₅₉X₁₁₆₀X₁₁₆₁X₁₁₆₂X₁₁₆₃X₁₁₆₄X₁₁₆₅X₁₁₆₆X₁₁₆₇X₁₁₆₈X₁₁₆₉X₁₁₇₀X₁₁₇₁X₁₁₇₂X₁₁₇₃X₁₁₇₄X₁₁₇₅X₁₁₇₆X₁₁₇₇X₁₁₇₈X₁₁₇₉X₁₁₈₀X₁₁₈₁X₁₁₈₂X₁₁₈₃X₁₁₈₄X₁₁₈₅X₁₁₈₆X₁₁₈₇X₁₁₈₈X₁₁₈₉X₁₁₉₀X₁₁₉₁X₁₁₉₂X₁₁₉₃X₁₁₉₄X₁₁₉₅X₁₁₉₆X₁₁₉₇X₁₁₉₈X₁₁₉₉X₁₂₀₀X₁₂₀₁X₁₂₀₂X₁₂₀₃X₁₂₀₄X₁₂₀₅X₁₂₀₆X₁₂₀₇X₁₂₀₈X₁₂₀₉X₁₂₁₀X₁₂₁₁X₁₂₁₂X₁₂₁₃X₁₂₁₄X₁₂₁₅X₁₂₁₆X₁₂₁₇X₁₂₁₈X₁₂₁₉X₁₂₂₀X₁₂₂₁X₁₂₂₂X₁₂₂₃X₁₂₂₄X₁₂₂₅X₁₂₂₆X₁₂₂₇X₁₂₂₈X₁₂₂₉X₁₂₃₀X₁₂₃₁X₁₂₃₂X₁₂₃₃X₁₂₃₄X₁₂₃₅X₁₂₃₆X₁₂₃₇X₁₂₃₈X₁₂₃₉X₁₂₄₀X₁₂₄₁X₁₂₄₂X₁₂₄₃X₁₂₄₄X₁₂₄₅X₁₂₄₆X₁₂₄₇X₁₂₄₈X₁₂₄₉X₁₂₅₀X₁₂₅₁X₁₂₅₂X₁₂₅₃X₁₂₅₄X₁₂₅₅X₁₂₅₆X₁₂₅₇X₁₂₅₈X₁₂₅₉X₁₂₆₀X<

KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KM food; technical material.
 XX
 OS Synthetic.
 XX
 PN WO9807745-A2.
 XX
 PD 26-FEB-1998.
 XX
 PF 21-AUG-1997; 97WO-US14779.
 XX
 PR 13-JAN-1997; 97US-0034949.
 PR 21-AUG-1996; 96US-0024754.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Erffle D, Fraser JR, Krieger TJ, Taylor R, West MH;
 DR WPI; 1998-169090/15.
 XX
 PT New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxyalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 XX
 PS Claim 11: Page 88; 129pp; English.
 XX
 CC AAY2549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VIII) containing up to 25 amino acids (aa): RxxXxxXB (I), BxxXxxXB
 CC (II), BxxXxxXxxXB (III), BxxXxxXxxBBn(AA)nMILBAGS (IV), BxxXxxXxxBB(AA)nM
 CC (V), LBnXxxXxxXxxnXR (VI), LKnXxxXxxXRR (VII) and BxxXxxXxxBB (VIII).
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 CC
 XX
 SQ Sequence 21 AA;
 Query Match 39.6%; Score 93.5; DB 19; Length 21;
 Best Local Similarity 63.0%; Pred. No. 1.3e-05;
 Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 QY 1 ILRNPMPWRRKHEAPEAPIMILKK 27
 Db 1 ILRNPMPWRRKHEAPEAPIMILKK 18
 XX
 RESULT 13
 AAY91798
 ID AAY91798 standard; Peptide; 21 AA.
 XX
 AC AAY91798;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Amino acid sequence of cationic peptide MBI 11B18CN.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;

KW leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KM multidrug resistance.
 XX
 OS Synthetic.
 XX
 PN WO9965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erffle D, Fraser JR, West MHP;
 DR WPI; 2000-223549/19.
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours
 PT
 XX
 PS Disclosure; Page 15; 94pp; English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 XX
 SQ Sequence 21 AA;
 Query Match 39.6%; Score 93.5; DB 21; Length 21;
 Best Local Similarity 63.0%; Pred. No. 1.3e-05;
 Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;
 QY 1 ILRNPMPWRRKHEAPEAPIMILKK 27
 Db 1 ILRNPMPWRRKHEAPEAPIMILKK 18
 XX
 RESULT 14
 AAW12873
 ID AAW12873 standard; peptide; 13 AA.
 XX
 AC AAW12873;
 XX
 DT 10-DEC-1997 (first entry)
 XX
 DE Antimicrobial cationic peptide CP-11.
 XX
 KW Bacterial; viral; antitumour; food; preservative; inhibitor; growth;
 KW bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;
 KW antiviral; Candida albicans; sterlant; Salmonella; Yersinia;
 KW Shigella.
 XX
 OS Synthetic.
 XX
 PN WO9708199-A2.
 XX
 PD 06-MAR-1997.
 XX
 PF 23-AUG-1996; 96WO-IB00996.
 XX
 PR 23-AUG-1995; 95US-0002667.
 XX

XX 21-AUG-1997; 97WO-US14779.
PF
XX

SEARCH COMPACTED
JOB TIME: 266 SEC

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:38:58 ; Search time 25.06 Seconds
(without alignments)
28.939 Million cell updates/sec

Title: SCHNIZ-444-MODSEQ2.PEP
Sequence: 236
1 ILRPMWPMWRKHEAPEPEPIMILKKMPMPWRRK 36

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	98.5	41.7	21	4	US-08-915-314-46 Sequence 46, Appl
2	94	39.8	20	4	US-08-915-314-47 Sequence 47, Appl
3	93.5	39.6	21	4	US-08-915-314-48 Sequence 48, Appl
4	91	38.6	13	4	US-08-915-314-30 Sequence 30, Appl
5	91	38.6	13	4	US-08-915-314-62 Sequence 62, Appl
6	91	38.6	13	4	US-08-915-314-63 Sequence 63, Appl
7	91	38.6	13	4	US-08-915-314-64 Sequence 64, Appl
8	91	38.6	13	4	US-08-915-314-57 Sequence 57, Appl
9	91	38.6	13	4	US-08-915-314-54 Sequence 54, Appl
10	87	36.9	12	4	US-08-915-314-52 Sequence 52, Appl
11	86	36.4	12	4	US-08-915-314-42 Sequence 42, Appl
12	86	36.4	12	4	US-08-915-314-74 Sequence 74, Appl
13	86	36.4	12	4	US-08-702-054B-5 Sequence 5, Appl
14	86	36.4	13	4	US-08-915-314-51 Sequence 51, Appl
15	86	36.4	13	4	US-08-915-314-58 Sequence 58, Appl
16	86	36.4	13	4	US-08-702-054B-34 Sequence 34, Appl
17	86	36.4	14	4	US-08-915-314-59 Sequence 59, Appl
18	86	36.4	15	4	US-08-702-054B-40 Sequence 40, Appl
19	86	36.4	20	4	US-08-915-314-55 Sequence 55, Appl
20	86	36.4	21	4	US-08-915-314-56 Sequence 56, Appl
21	85.5	36.2	16	4	US-08-702-054B-11 Sequence 11, Appl
22	85	36.0	12	4	US-08-915-314-69 Sequence 69, Appl
23	85	36.0	13	4	US-08-915-314-38 Sequence 38, Appl
24	85	36.0	13	4	US-08-915-314-45 Sequence 45, Appl
25	85	36.0	13	4	US-08-702-054B-1 Sequence 1, Appl
26	85	36.0	13	4	US-08-702-054B-17 Sequence 17, Appl
27	85	36.0	13	4	US-08-702-054B-32 Sequence 32, Appl

28	83.5	35.4	16	4	US-08-702-054B-38 Sequence 38, Appl
29	83	35.2	12	4	US-08-915-314-24 Sequence 24, Appl
30	83	35.2	12	4	US-08-915-314-40 Sequence 40, Appl
31	83	35.2	13	4	US-08-915-314-49 Sequence 49, Appl
32	83	35.2	13	4	US-08-915-314-50 Sequence 50, Appl
33	83	35.2	13	4	US-08-702-054B-30 Sequence 30, Appl
34	83	35.2	13	4	US-08-702-054B-31 Sequence 31, Appl
35	83	35.2	13	4	US-08-702-054B-35 Sequence 35, Appl
36	82	34.7	12	4	US-08-915-314-76 Sequence 76, Appl
37	82	34.7	13	4	US-08-915-314-25 Sequence 25, Appl
38	82	34.7	13	4	US-08-915-314-66 Sequence 66, Appl
39	82	34.7	13	4	US-08-915-314-67 Sequence 67, Appl
40	82	34.7	13	4	US-08-702-054B-33 Sequence 33, Appl
41	81	34.3	11	4	US-08-915-314-75 Sequence 75, Appl
42	81	34.3	12	4	US-08-915-314-77 Sequence 77, Appl
43	81	34.3	12	4	US-08-915-314-87 Sequence 87, Appl
44	81	34.3	15	4	US-08-702-054B-39 Sequence 39, Appl
45	80	33.9	12	4	US-08-915-314-78 Sequence 78, Appl

ALIGNMENTS

RESULT 1
US-08-915-314-46
Sequence 46, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOPLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
TELEPHONE: (206) 622-4900
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-46
Query Match 41.7%, Score 98.5, DB 4, Length 21;
Best Local Similarity 66.7%; Pred. No. 1.3e-06;
Matches 18, Conservative 0, Mismatches 0, Indels 9, Gaps 1;
CY 1 ILRPMWPMWRKHEAPEPEPIMILKK 27

DB 1 ILRWPMPWRRK-----MILK 18

RESULT 2

US-08-915-314-47

Sequence 47, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-47

Query Match

Best Local Similarity 39.8%; Score 94; DB 4; Length 20;

Matches 17; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILRWPMPWRRKHEAPEAPIMILK 27

DB 1 ILRWPMPWRRK-----MILK 17

RESULT 3

US-08-915-314-48

Sequence 48, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 48:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-48

Query Match

Best Local Similarity 39.6%; Score 93.5; DB 4; Length 21;

Matches 17; Conservative 0; Mismatches 1; Indels 9; Gaps 1;

OY 1 ILRWPMPWRRKHEAPEAPIMILK 27

DB 1 ILRWPMPWRRK-----MILK 18

RESULT 4

US-08-915-314-30

Sequence 30, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN

NUMBER OF SEQUENCES: 90

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 45:

SEQUENCE CHARACTERISTICS:

LENGTH: 21 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-48

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-30

Query Match
Best Local Similarity 38.6%; Score 91; DB 4; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPWPRRK 36
Db 1 ILKKPMPWPRRK 13

RESULT 5
US-08-915-314-62
Sequence 62, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="D-Form of Isoleucine"
US-08-915-314-62

Query Match
Best Local Similarity 38.6%; Score 91; DB 4; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPWPRRK 36
Db 1 ILKKPMPWPRRK 13

RESULT 6
US-08-915-314-63
Sequence 63, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note="D-Form of Lysine"
US-08-915-314-63

Query Match
Best Local Similarity 38.6%; Score 91; DB 4; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPWPRRK 36
Db 1 ILKKPMPWPRRK 13

RESULT 7
US-08-915-314-64
Sequence 64, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas

```

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note="D-Form of Isoleucine"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note="D-Form of Lysine"
US-08-915-314-64

Query Match          38.6%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 7.1e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ILKKPMPMPMRK 36
DB 1 ILKKPMPMPMRK 13

RESULT 8
US-08-915-314-57
Sequence 57, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
TELEFAX: (206) 622-4900
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-57

Query Match          38.6%; Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 7.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ILKKPMPMPMRK 36
DB 1 ILKKPMPMPMRK 13

RESULT 9
US-08-915-314-54
Sequence 54, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
TELEFAX: (206) 622-6031
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid

```

STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-54

Query Match 38.6%; Score 91; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 ILKKPMPWRRK 36
|||||
Db 1 ILKKPMPWRRK 13

RESULT 10
US-08-915-314-52
Sequence 52, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-52

Query Match 36.9%; Score 87; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.1e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 LKKPMPWRRK 36
|||||
Db 1 LKKPMPWRRK 12

RESULT 11
US-08-915-314-42
Sequence 42, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-42

Query Match 36.4%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRMPMPWRRK 12
|||||
Db 1 ILRMPMPWRRK 12

RESULT 12
US-08-915-314-74
Sequence 74, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION: Ph.D., Carol
NAME: No. 6180604tenburg 39,317
REGISTRATION NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-74

Query Match 36.4%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPMPRR 35
|||||
DB 1 ILKKPMPMPRR 12

RESULT 13
US-08-702-054B-5
Sequence 5, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J. W.
APPLICANT: Hancock, Robert E. W.
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-054B-5

Query Match 36.4%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 ILKKPMPMPRR 35
|||||
DB 1 ILKKPMPMPRR 12

RESULT 14
US-08-915-314-51
Sequence 51, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-6031
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 51:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-51

Query Match 36.4%; Score 86; DB 4; Length 13;
Best Local Similarity 41.9%; Pred. No. 3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 18; Gaps 1;

OY 6 WPPWRRKHEAPEAPITILKKPMPMPRRK 36
|||||
DB 1 WPPWRRKHEAPEAPITILKKPMPMPRRK 13

RESULT 15
US-08-915-314-58
Sequence 58, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
 TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
 NUMBER OF SEQUENCES: 90
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,314
 FILING DATE: 20-AUG-1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6180604tenburg Ph.D., Carol
 REGISTRATION NUMBER: 39,317
 REFERENCE/DOCKET NUMBER: 660081.405
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 58:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 13 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-915-314-58

Query Match 36.4%; Score 86; DB 4; Length 13;
 Best Local Similarity 100.0%; Pred. No. 3e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	24	ILKKPMPMPRR	35
DB	1	ILKKPMPMPRR	12

Search completed: June 21, 2001, 08:38:58
 Job time: 215 sec

GenCore version 4.5
Copyright (c) 1993 - 2000, CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:39:32 ; Search time 29.25 Seconds
(without alignments)
93.753 Million cell updates/sec

Title: SCHNIZ-444-MODSEQ2.PEP
Perfect score: 236
Sequence: 1 ILRPMWMPWRRRHEAPEAPDILMKWMPWRRK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:
1: pirl:*
2: pirl:*
3: pirl:*
4: pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78	33.1	144	1 JCI222	indollicidin precursor
2	63	26.7	192	2 H86543	hypothetical prote
3	63	26.7	192	2 D72081	conserved hypothet
4	61.5	26.1	2290	1 GNNYE	genome polypeptide
5	59.5	25.2	114	2 T36208	hypothetical prote
6	59	25.0	95	2 E86447	protein FSD14.5 [1
7	59	25.0	451	2 S30401	hypothetical prote
8	59	25.0	485	2 S74708	hypothetical prote
9	58	24.6	711	2 C40046	antibiotic transpo
10	57.5	24.4	376	2 T40591	hypothetical prote
11	55.5	23.5	983	1 VCLJVS	env polypeptide pr
12	55.5	23.5	983	1 E45390	hypothetical prote
13	55	23.3	352	2 S77448	hypothetical prote
14	55	23.3	452	2 T28094	hypothetical prote
15	55	23.3	715	2 B70741	probable moey prot
16	55	23.3	2292	2 S35961	capsid polypeptide
17	54.5	23.1	134	2 E72532	hypothetical prote
18	54	22.9	144	2 S35331	apidaecin 22 precu
19	54	22.9	424	2 T07742	omega-6 desaturase
20	54	22.9	469	2 B70607	probable cyts prot
21	54	22.9	1173	1 VG1HHC	E2 glycoprotein pr
22	53.5	22.7	538	2 B84759	hypothetical prote
23	53.5	22.7	990	2 E46335	env polypeptide pr
24	53.5	22.7	1113	2 JEO315	low-density lipopr
25	53	22.5	107	2 T35634	hypothetical prote
26	53	22.5	316	2 S43916	hypothetical prote
27	53	22.5	404	2 H69143	hypothetical prote
28	53	22.5	480	3 JC7552	coenzyme F420-redu
29	53	22.5	691	2 D71430	shb-like adapter p

30	53	22.5	949	2 E75352	glycine cleavage S
31	53	22.5	2292	1 GNNYED	genome polypeptide
32	53	22.5	2292	1 GNNYED	genome polypeptide
33	53	22.5	2292	2 S55401	capsid polypeptide
34	52.5	22.2	55	2 T11026	H+-transporting AT
35	52.5	22.2	295	2 C85783	probable excludin
36	52.5	22.2	295	2 E64933	hypothetical prote
37	52.5	22.2	443	2 T08136	probable Omega-6 d
38	52.5	22.2	448	2 D85362	hypothetical prote
39	52.5	22.2	449	2 C84618	hypothetical prote
40	52.5	22.2	471	2 T50016	transcription fact
41	52.5	22.2	486	2 T24334	hypothetical prote
42	52.5	22.2	490	2 T21365	hypothetical prote
43	52.5	22.2	518	2 T29589	hypothetical prote
44	52.5	22.2	621	2 S37664	peplomeric polypro
45	52.5	22.2	630	2 S37663	peplomeric polypro

ALIGNMENTS

RESULT 1

JCI222
Indollicidin precursor - bovine
N:Alternate names: antimicrobial peptide
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JCI222; A42387; S25664
R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.
Biochem. Biophys. Res. Commun. 187, 467-472, 1992
A:Title: cDNA cloning of the neutrophil bactericidal peptide indollicidin.
A:Reference number: JCI222; MUID:92392368
A:Accession: JCI222
A:Molecule type: mRNA
A:Residues: 1-144 <SAL>
A:Cross-references: EMBL:X67340; NID:9462; PIDN:CAA47755.1; PID:9463
A:Experimental source: bone marrow
R:Seisted, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.
J. Biol. Chem. 267, 4292-4295, 1992
A:Title: Indollicidin, a novel bactericidal tridecapeptide amide from neutrophils.
A:Reference number: A42387; MUID:92165771
A:Accession: A42387
A:Molecule type: protein
A:Residues: 131-143 <SEL>
A:Experimental source: neutrophils
A>Note: sequence extracted from NCBI backbone (NCBIP:83840)
C:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end
F:1-29/Domain: signal sequence #status predicted <SIG>
F:122-129/Domain: cystatin homology <CYTS>
F:30-130/Domain: propeptide #status predicted <PRO>
F:131-143/Product: indollicidin #status experimental <MAT>
F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from follow)

Query Match 33.1%; Score 78; DB 1; Length 144;
Best Local Similarity 55.6%; Pred. No. 0.0075;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 EAPDILMKWMPWRR 35
DB 126 ELQSVILPMKWMPWRR 143

RESULT 2
H86543
hypothetical protein Cpl0426 [imported] - Chlamydia pneumoniae (strain J138)
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: H86543
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A:Reference number: A86491; MUID:20330349
A:Accession: H86543
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-192 <STO>
C:Cross-references: GB:BA000008; NID:g8978798; PIDN:BAA98634.1; GSPDB:GN00142
C:Genetics:
A:Gene: CPJ0426

Query Match 26.7%; Score 63; DB 2; Length 192;
Best Local Similarity 42.3%; Pred. No. 0.74;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 4 WPMWPRRRKHEAPEAPIMILKKMP 29
 ||| | : | : | : | : | : |
DB 138 WPMLPKRQIEKLPRGEGICFLSAYP 163

RESULT 3
D12081
conserved hypothetical protein frameshifted CP0327 [Imported] - Chlamydia pneumoniae
N:Alternate names: hypothetical protein CT277 homolog
C:Species: chlamydia pneumoniae, Chlamydia pneumoniae
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: D12081; G81585
R:Kahnman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
Nature Genet. 21, 385-389, 1999
A>Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: D12081
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <ARN>
A:Cross-references: GB:AE001625; GB:AE001363; NID:g4376695; PIDN:AAD18570.1; PID:g437670
A:Experimental source: strain CML029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
N.C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A>Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A:Reference number: AB1500; MUID:20150255
A:Accession: G81589
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <REA>
A:Cross-references: GB:AE002195; GB:AE002161; NID:g7189246; PIDN:AAF38182.1; PID:g718925
A:Experimental source: strain AR39, HL cells
C:Genetics:

A:Gene: CPN0426; CP0327

Query Match 26.7%; Score 63; DB 2; Length 192;
Best Local Similarity 42.3%; Pred. No. 0.74;
Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 4 WPMWPRRRKHEAPEAPIMILKKMP 29
 ||| | : | : | : | : | : |
DB 138 WPMLPKRQIEKLPRGEGICFLSAYP 163

RESULT 4
GENNYE
genome polyprotein - encephalomyocarditis virus
N:Contents: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; core
EC 3.4.-.-); RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: encephalomyocarditis virus, EMCV
A>Note: host Homo sapiens (man)
C>Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
C:Accession: A03906; JN0383
R:Palmemery, A.C.; Kirby, E.M.; Janda, M.R.; Drake, N.L.; Duke, G.M.; Potratz, K.F.; Co
Nucleic Acids Res. 12, 2969-2985, 1984
A>Title: The nucleotide and deduced amino acid sequences of the encephalomyocarditis vir
A:Reference number: A03906; MUID:84169586

A:Accession: A03906
A:Molecule type: genomic RNA
A.Residues: 1-2290 <PAL>
A.Cross-references: GB:X00463; NID:g61034; PIDN:CAA25152_1; PTD:g61035
R.Petrov, N.A.; Chizhikov, V.E.; Blinov, V.M.; Karlinov, V.A.; Mkrtyukov, N.N.; Gutor
Bicoxy. Khim. 10, 274-279, 1984
A.Title: Nucleotide sequence of the 3'-terminus of encephalomyocarditis virus RNA.
A.Feature number: JN0383; MUID:85022788
A.Accession: JN0383
A.Molecule type: genomic RNA
A.Residues: 1337-1396, 'L', 1398-1517, 'A', 1519-1556, 'E', 1538-1556, 'S', 1558-1611, 'T', 1611-
A.Cross-references: GB:M54933
A.Note: The authors translated the codon CAU for residue 713 as Thr and AAC for resid
C.Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleotidylyl
F:1-67/Domain: leader peptide #status: predicted <LDP>
F:168-136/Product: coat protein VP4 #status: predicted <VP4>
F:137-391/Product: coat protein VP2 #status: predicted <VP2>
F:392-622/Product: coat protein VP3 #status: predicted <VP3>
F:623-910/Product: coat protein VP1 #status: predicted <VP1>
F:911-1056/Product: core protein p2-A #status: predicted <p2A>
F:1057-1192/Product: core protein p2-B #status: predicted <p2B>
F:1193-1517/Product: core protein p2-C #status: predicted <p2C>
F:1518-1605/Product: core protein p3-A #status: predicted <p3A>
F:1606-1625/Product: genome-linked protein VPg #status: predicted <VPg>
F:1626-1830/Product: proteinase #status: predicted <PTS>
F:1831-2290/Product: RNA-directed RNA polymerase #status: predicted <RDP>

```

Query Match      26.1%  Score 61.5;  DB 1;  Length 2290;
Best Local Similarity 27.3%  Pred. No. 15;
Matches 12;  Conservative 6;  Mismatches 9;  Indels 17;  Gaps 1;

OY      3  RPPWPMWRKRHEAPEAPEPI-----MILKKMP 29
      | | | | | : : | | | : : | | |
Db      965  RAPNPNMKNKTQAVLARAEPCRYTMDIYYKRVPRFLPLVQKEMP 1008

RESULT      5
T36208
hypothetical protein SCE36.09 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36208
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21601
A:Accession: T36208
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-114 <OL>
A:Cross-references: EMBL:AL049763; PIDN:CA842078.1; GSPDB:GN00070; SCODEB:SCE36.09
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCODEB:SCE36.09

Query Match      25.2%  Score 59.5;  DB 2;  Length 114;
Best Local Similarity 50.0%  Pred. No. 1.2;
Matches 10;  Conservative 2;  Mismatches 7;  Indels 1;  Gaps 1;

OY      17  PEAEPIILKKM-PWMPWRR 35
      | | | : : | | | | |
Db      93  PEFAPADARWRMRPWPWRR 112

RESULT      6
E86447
protein FSD14.5 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: E86447
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alton

```

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hutzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzall,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: E86447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-95 <STO>
A:Cross-references: GB:AE005172; NID:98920603; PIDN:AAF81325.1; GSPDB:GN00141
C:Genetics:
A:Gene: F5D14.5
A:Map position: 1

Query Match 25.0%; Score 59; DB 2; Length 95;
Best Local Similarity 21.3%; Pred. No. 1.1;
Matches 10; Conservative 5; Mismatches 4; Indels 28; Gaps 3;

QY 4 WP-----WMPRRKHEAPEAPIMILK---KMPMPW 33
DB 46 MPVVVVVAGVGGRMMNMV-----PVLVITDVGVGEMSWMW 81

RESULT 7
S30401
hypothetical protein 2 - Streptomyces clavuligerus plasmid pSCL
C:Species: Streptomyces clavuligerus
C:Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 22-Oct-1999
C:Accession: S30401
R:Wu, X.; Roy, K.L.
J. Bacteriol. 175, 37-52, 1993
A:Title: Complete nucleotide sequence of a linear plasmid from Streptomyces clavuligerus
A:Reference number: S30400; MUID:93106972
A:Accession: S30401
A:Molecule type: DNA
A:Residues: 1-451 <MX>
A:Cross-references: EMBL:X54107; NID:948756; PIDN:CAA38041.1; PID:9581632
C:Genetics:
A:Genome: plasmid pSCL
A:Start codon: GTG

Query Match 25.0%; Score 59; DB 2; Length 451;
Best Local Similarity 25.0%; Pred. No. 5.6;
Matches 15; Conservative 3; Mismatches 4; Indels 38; Gaps 4;

QY 2 LRMP-----WMPRRKHEAPEAPIMILKMP-----WMPW 33
DB 362 LRMPFEGFGLSPATMGAEELRLMDLPMW-----GPRSE-----RMPAFCGCTMPW 411

RESULT 8
S74708
hypothetical protein slr1306 - Synecocystis sp. (strain PCC 6803)
C:Species: Synecocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S74708
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis
S:Reference number: S74322; MUID:97061201
A:Accession: S74708
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-485 <KAN>
A:Cross-references: EMBL:D90901; GB:AB001339; NID:91651897; PIDN:BA16859.1; PID:d101
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 25.0%; Score 59; DB 2; Length 485;
Best Local Similarity 29.3%; Pred. No. 6.1;
Matches 12; Conservative 6; Mismatches 5; Indels 18; Gaps 2;

QY 5 PW-----WMPRRKHEAPEAPIM---ILK 27
DB 41 PWDQMGIALSLGLVIWPMRRHRHAPPEOKMLPEVLOK 81

RESULT 9
C40046
antibiotic transport-associated protein actII-3 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Oct-1999
C:Accession: C40046
R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.
Cell 66, 769-780, 1991
A:Title: The act cluster contains regulatory and antibiotic export genes, direct tar
A:Reference number: A40046; MUID:91347376
A:Accession: C40046
A:Molecule type: DNA
A:Residues: 1-711 <FER>
A:Cross-references: GB:M64683; NID:9153143; PIDN:AAA26691.1; PID:9153146

Query Match 24.6%; Score 58; DB 2; Length 711;
Best Local Similarity 56.2%; Pred. No. 12;
Matches 9; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 3 RMPMPMPRRKHEAPE 18
DB 335 RMPMPMPRRKHGTEPD 350

RESULT 10
T40591
hypothetical protein SPBC646.15c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 04-Mar-2000
C:Accession: T40591
R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, January 1999
A:Reference number: Z21938
A:Accession: T40591
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-376 <SEE>
A:Cross-references: EMBL:AL035216; PIDN:CAA22819.1; GSPDB:GN00067; SPDB:SPBC646.15c
A:Experimental source: strain 972n-; cosmid c646
C:Genetics:
A:Gene: SPDB:SPBC646.15c
A:Map position: 2
A:Inserts: 49/1; 126/2; 312/2; 350/1
C:Superfamily: Schizosaccharomyces pombe hypothetical protein SPBC646.15c

Query Match 24.4%; Score 57.5; DB 2; Length 376;
Best Local Similarity 37.0%; Pred. No. 7.2;
Matches 10; Conservative 3; Mismatches 7; Indels 7; Gaps 1;

QY 7 WMPRRKHEAPEAPIMILKMPMPW 33
DB 236 WMPRRKOKSSS-----LKVPMGPW 255

RESULT 11
VCLJVS
env polyprotein precursor - Meedi/Visna virus (strain 1514)

N:Alternate names: coat polyprotein
 C:Species: Maedi/Visna virus
 A:Note: host Homo sapiens (man)
 C>Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 16-Feb-1997
 C:Accession: A03977
 R:Sonigo, P.; Alloum, M.; Staskus, K.; Klatzmann, D.; Cole, S.; Danos, O.; Retzel, E.; T
 Cell 42, 359-382, 1985
 A:Title: Nucleotide sequence of the visna lentivirus: relationship to the AIDS virus.
 A:Reference number: A90869; MUID:65254938
 A:Accession: A03977
 A:Molecule type: DNA
 A:Residues: 1-982 <SON>
 C:Genetics:
 A:Gene: env
 C:Superfamily: visna lentivirus type E retrovirus env polyprotein
 C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
 F:1100/Domain: signal sequence #status predicted <SIG>
 F:101-556/Product: exterior membrane glycoprotein #status predicted <EXT>
 F:657-982/Product: transmembrane glycoprotein #status predicted <TM>
 F:140,161,206,258,298,364,370,381,387,403,414,435,439,470,475,481,491,501,515,527,537,54
 F:697,764,771,787,821/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 23.5%; Score 55.5; DB 1; Length 982;
 Best Local Similarity 34.3%; Pred. No. 35;
 Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

OY 1 ILRPMW-----PWRKHEAPEAPIMILKKMPW 30
 Db 911 IWRATWAMKTSFWRHNRMTPTTLPLIVIMQW 945

RESULT 12
 E45390
 env polyprotein precursor - Maedi/Visna virus (strain KV1772) (provirus)
 N:Alternate names: coat polyprotein
 N:Contains: exterior membrane glycoprotein; transmembrane glycoprotein
 C:Species: Maedi/Visna virus
 A:Note: host Homo sapiens (man)
 C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C:Accession: E45390
 R:Andersson, O.S.; Elser, J.E.; Tobin, G.J.; Greenwood, J.D.; Gonda, M.A.; Georgsson, G.
 J.W.; Petrusson, G.
 Virology 193, 89-105, 1993
 A:Title: Nucleotide sequence and biological properties of a pathogenic proviral molecule
 A:Reference number: A45390; MUID:93174981
 A:Accession: E45390
 A:Molecule type: DNA
 A:Residues: 1-983 <AND>
 A:Cross-references: GB:S55323; NID:g265825; PIDN:AB25463.1; PID:g265830
 C:Genetics:
 A:Gene: env
 C:Superfamily: visna lentivirus type E retrovirus env polyprotein
 C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
 F:1100/Domain: signal sequence #status predicted <SIG>
 F:101-556/Product: exterior membrane glycoprotein #status predicted <EXT>
 F:657-983/Product: transmembrane glycoprotein #status predicted <TM>
 F:836-852/Domain: transmembrane #status predicted <TM>
 F:140,161,206,258,298,364,381,387,403,414,435,439,470,475,481,491,501,515,527,537,542,54

Query Match 23.5%; Score 55.5; DB 1; Length 983;
 Best Local Similarity 34.3%; Pred. No. 35;
 Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

OY 1 ILRPMW-----PWRKHEAPEAPIMILKKMPW 30
 Db 912 IWRATWAMKTSFWRHNRMTPTTLPLIVIMQW 946

RESULT 13
 S77448
 hypothetical protein s11080 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S77448
 R:Kaneh, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 O.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-136, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
 S.
 A:Reference number: S74322; MUID:97061201
 A:Accession: S77448
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-352 <KAN>
 A:Cross-references: EMBL:D90905; GB:AB001339; NID:91652360; PIDN:BA17295.1; PID:d101
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 23.3%; Score 55; DB 2; Length 352;
 Best Local Similarity 52.6%; Pred. No. 14;
 Matches 10; Conservative 1; Mismatches 6; Indels 2; Gaps 2;

OY 17 PEAPIMI-LKKMP-WMPW 33
 Db 47 PEGPPIYIGSNMAGWMPW 65

RESULT 14
 T28094
 hypothetical protein ZK899.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Mar-2000
 C:Accession: T28094
 R:Kershaw, J.
 submitted to the EMBL Data Library, September 1994
 A:Reference number: Z20468
 A:Accession: T28094
 A:Status: preliminary; translated from GB/EMBL/DDB
 A:Molecule type: DNA
 A:Residues: 1-452 <WIL>
 A:Cross-references: EMBL:Z37140; PIDN:CAA85502.1; GSPDB:GN00028; CESP:ZK899.2
 C:Experimental source: clone ZK899
 C:Genetics:
 A:Gene: CESP:ZK899.2
 A:Map position: X
 A:Insertions: 34/3; 143/2; 227/2; 262/3; 380/3
 C:Superfamily: Caenorhabditis elegans hypothetical protein ZK899.2

Query Match 23.3%; Score 55; DB 2; Length 452;
 Best Local Similarity 32.1%; Pred. No. 18;
 Matches 9; Conservative 3; Mismatches 8; Indels 8; Gaps 2;

OY 6 WMPRRKHEAPEAPIMILKKMPWMPW 33
 Db 175 WMTW----HDDPN-----IDRKNMWTW 194

RESULT 15
 B70741
 probable moxy protein - Mycobacterium tuberculosis (strain H37Rv)
 C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: B70741
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon
 ; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
 A:Reference number: A70500; MUID:98295987
 A:Accession: B70741
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

Query Match	23.38;	Score 55;	DB 2;	Length 715;
Best Local Similarity	38.18;	Pred. No. 29;		
Matches	8;	Conservative	7;	Mismatches 4;
				Indels 2;
				Gaps 1;

```
QY      17 PEAEPIMI--LKKWPWPWRR      35
          | : | : : | : | : | | |
Db      53 PDPDDLLEAKRWAYYPWR      73
```

Search completed: June 21, 2001, 08:39:32
Job time: 158 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:44:55 ; Search time 17.73 Seconds
(without alignments)
69.554 Million cell updates/sec

Title: SCINIZ-444-MODSEQ2.PEP
 Perfect score: 236
 Sequence: 1 ILRPMPMPRRKHEAPEAEPIMLKKMPMPMPRRK 36

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

```
Searched:      93435 seqs, 34255486 residues
Total number of hits satisfying chosen parameters:  93435
```

```

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

```

```

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

Database : SwissProt_39: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	78	33.1	144	1	INDC_BOVIN	P33046 bos taurus
2	61.5	26.1	2290	1	POLC_EMCV	P03304 encephalomy
3	58	24.6	711	1	MLLA_STRIC	O53902 streptomyce
4	55.5	23.5	55	1	ATP8_AANPL	P50655 anas platyr
5	55.5	23.5	982	1	ENV_VILV	P03379 visna lentii
6	55.5	23.5	983	1	ENV_VILV	P35954 visna lentii
7	55.5	23.5	991	1	ENV_VILV2	P23423 visna lentii
8	55	23.3	715	1	YD55_MCTV2	Q11025 mycobacteri
9	54	22.9	144	1	AP22_APTME	P35581 apis mellif
10	54	22.9	424	1	PD6C_SOYBN	P48628 glycyne max
11	54	22.9	469	1	SYCL_MCTV2	P56862 mycobacteri
12	54	22.9	1173	1	VG12_CV422	P15423 human coron
13	53.5	22.7	990	1	ENV_OMVVS	P16899 ovine lentii
14	53	22.5	314	1	YMA3_BACST	O45633 bacillus st
15	53	22.5	2292	1	POLG_EMCVB	P17593 encephalomy
16	53	22.5	2292	1	POLG_EMCVD	P17594 encephalomy
17	52.5	22.2	55	1	ATP8_AYRKM	O94K25 aythya amer
18	52.5	22.2	257	1	E434_ADESO	O64865 human adeno
19	52.5	22.2	418	1	PD6C_ARATH	P46312 arabidopsis
20	52.5	22.2	443	1	ED6C_BRANA	P46327 brassica na
21	52.5	22.2	691	1	YHUG_ECOLI	P37645 escherichia
22	52.5	22.2	824	1	TGLK_RAT	P22606 rattus norv
23	52.5	22.2	989	1	ENV_VILV1	P23442 visna lentii
24	52.5	22.2	1154	1	VG12_IBVD2	P11223 avian infec
25	52.5	22.2	1162	1	VG12_IBVB	P11223 avian infec
26	52.5	22.2	1162	1	VG12_IBVM	P12650 avian infec
27	52.5	22.2	1162	1	VG12_IBVM	P12651 avian infec
28	52.5	22.2	1163	1	VG12_IBV6	P05135 avian infec
29	52	22.0	68	1	Y121_BP74	Q02405 bacteriophan
30	51.5	21.8	556	1	MEND_ECOLI	P17409 menaquinol
31	51	21.6	126	1	YD43_MCTV	Q11013 mycobacteri
32	51	21.6	196	1	YAO5_SCHPO	O09677 schizosacch
33	51	21.6	413	1	YBHO_ECOLI	P75771 escherichia

34	51	21.6	1003	1	POL_HV1A2	P03369	human	immun
35	51	21.6	1003	1	POL_HV107	P20892	human	immun
36	51	21.6	1016	1	PDQG_EC0L1	P33176	eschertichia	
37	50.5	21.4	887	1	UFO_HUMAN	P30530	homo sapien	
38	50	21.2	79	1	YAAO_BACCU	P37509	bacillus su	
39	50	21.2	157	1	NUQC_SYNY3	P19125	synecocyst	
40	50	21.2	283	1	AP73_APIME	Q06602	apis mellif	
41	50	21.2	465	1	TPSN_MOUSE	Q91733	mus musculu	
42	50	21.2	984	1	SN13_MOUSE	Q04891	mus musculu	
43	49.5	21.0	287	1	YHJY_SALT1	P50335	salmoneila	
44	49.5	21.0	396	1	O45B_DROME	Q9Y889	drosophila	
45	49.5	21.0	478	1	GATB_AQUAE	O66766	aquifex aeo	

```

RESULT 1
INDC_BOVIN STANDARD; PRT; 144 AA.
ID INDC_BOVIN STANDARD; PRT; 144 AA.
DT P3046; 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-OCT-1993 (Rel. 27, Last annotation update)
DE INDOLICIDIN PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId:9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE:92392368; PubMed:1520337;
RA del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;
RT "CDNA cloning of the neutrophil bactericidal peptide indolecidin."
RL Biochem. Biophys. Res. Commun. 187:467-472(1992).
RN [2]
RP SEQUENCE OF 131-143.
RC TISSUE=Neutrophils;
RX MEDLINE:92165771; PubMed:1537821;
RA Seasted M.E., Novotny M.J., Morris W.L., Tang Y.-O., Smith W.,
RA Cullor J.S.;
RT "Indolecidin, a novel bactericidal tridecapeptide amide from
RT neutrophils."
RL J. Biol. Chem. 267:4292-4295(1992).
CC -I- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST
CC STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.
CC -I- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.
CC -I- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.
CC -I- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: X67340; CAA47755.1;
DR PIR: JCI222; JCI222.
DR PIR: A42387; A42387.
DR InterPro: IPR001894;
DR Pfam: PF00666; Cathelicidins: 1.
DR PROSITE: PS00946; CATHELICIDINS_1; 1.
DR PROSITE: PS00947; CATHELICIDINS_2; 1.
KW Antibiotic; Amidation; Signal.
FT SIGNAL 1 29 POTENTIAL.
FT PROPEP 30 130
FT PEPTIDE 131 143
FT MOD_RES 30 30
FT INDOLECIDINE
FT PYROGLUTAMINE
FT CARBOXYLIC ACID (BY
FT SIMILARITY).

```

FT DISULEID 85 96 BY SIMILARITY.
 FT DISULEID 107 124 BY SIMILARITY.
 FT MOD_RES 143 143 AMIDATION (G-144 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 144 AA; 14479 MW; E3B1CB8E55C09911 CRC64;

Query Match 33.1%; Score 78; DB 1; Length 144;
 Best Local Similarity 55.6%; Pred. No. 0.003;
 Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 18 EAEPIMLKRWPMWPMRR 35
 Db 126 ELQSVILPMKRWPMWPMRR 143

RESULT 2
 ID POLG_EMCV STANDARD; PRT; 2290 AA.
 AC P03304;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEIN VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C, P3A; GENOME-LINKED PROTEIN VP6; PICORNAIN 3C
 DE (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D
 DE (EC 2.7.7.48)].
 OS Encephalomyocarditis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirus.
 NC NCBI_TaxID=12104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=84169586; Pubmed=6324136;
 RA Palmenberg A.C., Kirby E.M., Janda M.R., Drake N.L., Duke G.M.,
 RA Politz K.F., Collect M.S.;
 RT "The nucleotide and deduced amino acid sequences of the
 RT Encephalomyocarditis viral polyprotein coding region.";
 RL Nucleic Acids Res. 12:2969-2985(1984).
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC Q/Q SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X00463; CAA25152.1; -
 DR PIR: A03906; GNNE.
 DR HSSP: P12296; IMEC.
 DR MEROPS: C03.009; -
 DR InterPro: IPR000605; -
 DR InterPro: IPR001205; -
 DR InterPro: IPR001676; -
 DR Pfam: PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam: PF00910; RNA_helicase; 1.
 KM Polypeptide; Coat protein; Core protein; Transferase;
 KM RNA-directed RNA polymerase; Hydroxylase; Thiol protease; Myristate.
 FT PROPEP 1 67 LEADER PEPTIDE.
 FT CHAIN 68 136 COAT PROTEIN VP4 (RHO).
 FT CHAIN 137 391 COAT PROTEIN VP2 (BETA).
 FT CHAIN 392 622 COAT PROTEIN VP3 (GAMMA).
 FT CHAIN 623 910 COAT PROTEIN VP1 (ALPHA).
 FT CHAIN 911 1056 CORE PROTEIN P2A (G).

FT CHAIN 1057 1192 CORE PROTEIN P2B (I).
 FT CHAIN 1193 1517 CORE PROTEIN P2C (F).
 FT CHAIN 1518 1605 CORE PROTEIN P3A.
 FT CHAIN 1606 1625 GENOME-LINKED PROTEIN VP6 (H).
 FT CHAIN 1626 1830 PICORNAIN 3C (P22).
 FT CHAIN 1831 2290 RNA-DIRECTED RNA POLYMERASE P3D (E).
 FT LIPID 68 68 MYRISTATE (BY SIMILARITY).
 FT ACT_SITE 1784 1784 PROTEASE (POTENTIAL).
 FT ACT_SITE 1802 1802 PROTEASE (POTENTIAL).
 SQ SEQUENCE 2290 AA; 255756 MW; 26BC81B7CFC68CB5 CRC64;

Query Match 26.1%; Score 61.5; DB 1; Length 2290;
 Best Local Similarity 27.3%; Pred. No. 5.6;
 Matches 12; Conservative 6; Mismatches 9; Indels 17; Gaps 1;

QY 3 RMPWPMRRKHEAPEAPEPI-----MILKWP 29
 Db 965 RAPWPMKNTYQAVLRAPPCRVMTDIYKRYRPRPLPYQKEWP 1008

RESULT 3
 ID MMLA_STRCO STANDARD; PRT; 711 AA.
 AC O53902;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE POTATIVE MEMBRANE PROTEIN ACTII-3.
 DE ACTII-3.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
 NC NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91347376; Pubmed=1878971;
 RA Fernandez-Moreno M.A., Caballero J.L., Hopwood D.A., Malpartida F.;
 RT "The act cluster contains regulatory and antibiotic export genes,
 RT direct targets for translational control by the bldA tRNA gene of
 RT Streptomyces.";
 RL Cell 66:769-780(1991).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE MMLP FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M64683; AAA26691.1; -
 DR PROSITE: PS50156; SSD; 2.
 KM Hypothetical protein; Transmembrane.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 235 255 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 FT TRANSMEM 313 333 POTENTIAL.
 FT TRANSMEM 369 389 POTENTIAL.
 FT TRANSMEM 516 536 POTENTIAL.
 FT TRANSMEM 540 560 POTENTIAL.
 FT TRANSMEM 573 593 POTENTIAL.
 FT TRANSMEM 623 643 POTENTIAL.
 FT TRANSMEM 645 665 POTENTIAL.
 SQ SEQUENCE 711 AA; 74862 MW; A5466BEDABEDB16 CRC64;

Query Match 24.6%; Score 58; DB 1; Length 711;
 Best Local Similarity 56.2%; Pred. No. 4.8;

DR	Pfam: PF00517; Gp41; 1.	Polypeptide; Transmembrane.	LEADER PEPTIDE.	EXTERIOR MEMBRANE GLYCOPROTEIN.	POTENTIAL.
DR	Glycoprotein; Coat protein; Peptide	1	100		
FM	CHAIN	101	664		
FM	CHAIN	665	991		
FM	TRANSMEM	844	860		
FM	CARBOHYD	140	140		
FM	CARBOHYD	151	151		
FM	CARBOHYD	268	268		
FM	CARBOHYD	258	258		
FM	CARBOHYD	298	298		
FM	CARBOHYD	354	354		
FM	CARBOHYD	361	381		
FM	CARBOHYD	387	387		
FM	CARBOHYD	403	403		
FM	CARBOHYD	435	435		
FM	CARBOHYD	439	439		
FM	CARBOHYD	470	470		
FM	CARBOHYD	475	475		
FM	CARBOHYD	481	481		
FM	CARBOHYD	491	491		
FM	CARBOHYD	501	501		
FM	CARBOHYD	515	515		
FM	CARBOHYD	527	527		
FM	CARBOHYD	537	537		
FM	CARBOHYD	542	542		
FM	CARBOHYD	547	547		
FM	CARBOHYD	552	552		
FM	CARBOHYD	557	557		
FM	CARBOHYD	562	562		
FM	CARBOHYD	567	567		
FM	CARBOHYD	572	572		
FM	CARBOHYD	577	577		
FM	CARBOHYD	582	582		
FM	CARBOHYD	587	587		
FM	CARBOHYD	592	592		
FM	CARBOHYD	597	597		
FM	CARBOHYD	602	602		
FM	CARBOHYD	607	607		
FM	CARBOHYD	612	612		
FM	CARBOHYD	617	617		
FM	CARBOHYD	622	622		
FM	CARBOHYD	627	627		
FM	CARBOHYD	632	632		
FM	CARBOHYD	637	637		
FM	CARBOHYD	642	642		
FM	CARBOHYD	647	647		
FM	CARBOHYD	652	652		
FM	CARBOHYD	657	657		
FM	CARBOHYD	662	662		
FM	CARBOHYD	667	667		
FM	CARBOHYD	672	672		
FM	CARBOHYD	677	677		
FM	CARBOHYD	682	682		
FM	CARBOHYD	687	687		
FM	CARBOHYD	692	692		
FM	CARBOHYD	697	697		
FM	CARBOHYD	702	702		
FM	CARBOHYD	707	707		
FM	CARBOHYD	712	712		
FM	CARBOHYD	717	717		
FM	CARBOHYD	722	722		
FM	CARBOHYD	727	727		
FM	CARBOHYD	732	732		
FM	CARBOHYD	737	737		
FM	CARBOHYD	742	742		
FM	CARBOHYD	747	747		
FM	CARBOHYD	752	752		
FM	CARBOHYD	757	757		
FM	CARBOHYD	762	762		
FM	CARBOHYD	767	767		
FM	CARBOHYD	772	772		
FM	CARBOHYD	777	777		
FM	CARBOHYD	782	782		
FM	CARBOHYD	787	787		
FM	CARBOHYD	792	792		
FM	CARBOHYD	797	797		
FM	CARBOHYD	802	802		
FM	CARBOHYD	807	807		
FM	CARBOHYD	812	812		
FM	CARBOHYD	817	817		
FM	CARBOHYD	822	822		
FM	CARBOHYD	827	827		
FM	CARBOHYD	832	832		
FM	CARBOHYD	837	837		
FM	CARBOHYD	842	842		
FM	CARBOHYD	847	847		
FM	CARBOHYD	852	852		
FM	CARBOHYD	857	857		
FM	CARBOHYD	862	862		
FM	CARBOHYD	867	867		

```

FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 705 705 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 796 796 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 830 830 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 17 17 E -> D (IN AA17532).
SQ SEQUENCE 991 AA: 115016 MW: D8920FE4A8A55A7 CRC64:

Query Match 23.5%: Score 55.5; DB 1; Length 991;
Best Local Similarity 34.3%; Pred. No. 14;
Matches 12; Conservative 1; Mismatches 17; Indels 5; Gaps 1;

OY 1 ILRPMW-----PMRRKHEAPEPIMILKKMPW 30
DB 920 IWRATWMAKTSFMRHNRMTPTLPLPLVIMQW 954

RESULT 8
YD55_MYCTU STANDARD: PRT: 715 AA.
ID YD55_MYCTU
AC Q11025;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE HYPOTHETICAL 78.2 KDA PROTEIN RV1355C.
GN RV1355C OR MTCY02B10.19C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: Z75555; CA99988.1;
CC Tuberculist; RV1355C;
CC InterPro: IPR000594;
CC Pfam: PF00899; Th1F family; 1.
CC Hypothetical protein.
CC SEQUENCE 715 AA: 78181 MW: 455495248A56041C CRC64:

Query Match 23.3%; Score 55; DB 1; Length 715;
Best Local Similarity 38.1%; Pred. No. 11;
Matches 8; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

OY 17 PEAEPIMT--LKKWPMPPRR 35
DB 53 POPDPDLLEAKRMAYVPMRR 73

```

```

RESULT 9
AP22_APIME STANDARD: PRT: 144 AA.
ID AP22_APIME
AC P35581; P11525; P11526;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE APIDAEIN PRECURSOR, TYPE 22.
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Apoidea; Apidae; Apis.
OX NCBI_TaxID=7460;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=93223697; PubMed=8467807;
RA Casteels J., Ampe C., Jacobs F., Vaeck M., Tempst P.;
RA "Apidaein multipetide precursor structure: a putative mechanism for
RT amplification of the insect antibacterial response."
RN EMBO J. 12:1569-1578(1993).
RN (2)
RP SEQUENCE (APIDAEIN IA/IB).
RC TISSUE=Hemolymph;
RC MEDLINE=90003446; PubMed=2676519;
RA Casteels P., Ampe C., Jacobs F., Vaeck M., Tempst P.;
RA "Apidaeins: antibacterial peptides from honeybees."
RL EMBO J. 8:2387-2391(1989).
CC CC -1- FUNCTION: APIDAEIN HAVE BACTERICIDAL ACTIVITY; PREDOMINANTLY
CC AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL
CC PROPAGATION.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X72576; CA51168.1;
CC PIR: S05383; S05383.
CC PIR: S06675; S06675.
CC PIR: S35331; S35331.
CC InterPro: IPR001979;
CC Pfam: PF00807; Apidaein; 4.
CC Insect immunity; Antibiotic; Hemolymph; Signal; Multigene family;
KW Cleavage on pair of basic residues; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 35 42 APIDAEIN IB.
FT PEPTIDE 43 60 APIDAEIN IB.
FT PROPEP 63 70 APIDAEIN IB.
FT PEPTIDE 71 88 APIDAEIN IB.
FT PROPEP 91 98 APIDAEIN IB.
FT PEPTIDE 99 116 APIDAEIN IB.
FT PROPEP 119 126 APIDAEIN IB.
FT PEPTIDE 127 144 APIDAEIN IA.
SQ SEQUENCE 144 AA: 16539 MW: 6FMA1AD7ACB7108D CRC64:

Query Match 22.9%; Score 54; DB 1; Length 144;
Best Local Similarity 64.7%; Pred. No. 3.1;
Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 5 PMWPMRRKHEAPEAPEP 21
DB 26 PTRPRLRREAEPEAPEP 42

RESULT 10
FD6C_SOYBN STANDARD: PRT: 424 AA.
ID FD6C_SOYBN
AC P48628;

```

```

DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE OMEGA-6 FATTY ACID DESATURASE, CHLOROPLAST PRECURSOR (EC 1.14.99.-).
OS Glycine max (soybean).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC Fabales; Fabaceae; Papilionoideae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Seed;
RX MEDLINE=9445008; PubMed=8066133;
RA Hitz W.D., Carlson T.J., Booth J.R. Jr., Kinney A.J., Stecca K.L.,
RA Yadav N.S.;
RT Cloning of a higher-plant plastid omega-6 fatty acid desaturase cDNA
RT and its expression in a cyanobacterium."
RL Plant Physiol. 105:635-641(1994).
CC -1- FUNCTION: CHLOROPLAST OMEGA-6 FATTY ACID DESATURASE INTRODUCES
CC THE SECOND DOUBLEBOND IN THE BIOSYNTHESIS OF 16:3 AND 18:3 FATTY
CC ACIDS, IMPORTANT CONSTITUENTS OF PLANT MEMBRANES. IT IS THOUGHT
CC TO USE FERREDOXIN AS AN ELECTRON DONOR AND TO ACT ON FATTY ACIDS
CC ESTERIFIED TO GALACTOLIPIDS, SULFOLIPIDS AND PHOSPHATIDYLGLYCEROL.
CC -1- PATHWAY: POLYUNSATURATED FATTY ACID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST, MEMBRANE-BOUND (PROBABLE).
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: TO OTHER PLANT CHLOROPLAST OMEGA-6 FATTY ACID
CC DESATURASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L29215; AAA50158.1; -
DR InterPro: IPR001225; -
DR Pfam: PF00487; FA_desaturase; 1.
KW Oxidoreductase; Fatty acid biosynthesis; Chloroplast; Membrane;
KW Transf. Peptide.
FT TRANSIT 1 63 CHLOROPLAST (BY SIMILARITY).
FT CHAIN 64 424 OMEGA-6 FATTY ACID DESATURASE.
FT DOMAIN 165 169 HISTIDINE BOX 1.
FT DOMAIN 201 205 HISTIDINE BOX 2.
FT DOMAIN 361 365 HISTIDINE BOX 3.
SO SEQUENCE 424 AA; 49641 MW; 4F3DF52D4B1A2009 CRC64;

Query Match 22.9%; Score 54; DB 1; Length 424;
Best Local Similarity 35.3%; Pred. No. 9;
Matches 12; Conservative 7; Mismatches 13; Indels 2; Gaps 2;

OY 2 LRPWMPRRKHEAPPEAPPEPMLIKKMPMP-WR 34
DB 190 LIRPEPWRKHDRH-HAKTNMLREDTPAMPVWK 222

RESULT 11
SYCL_MYCTU STANDARD; PRT; 469 AA.
ID SYCL_MYCTU
AC P96862;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYSTEINYL-TRNA SYNTHETASE 1 (EC 6.1.1.16) (CYSTEINE--TRNA LIGASE 1)
DE (CYRS 1).
GN CYSS1 OR CYSS OR RV3580C OR MTCY06G11.27C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.

```

```

OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
CC -1- CATALYTIC ACTIVITY: ATP + L-CYSTEINE + TRNA(CYS) -> AMP +
CC PYROPHOSPHATE + L-CYSTEINYL-TRNA(CYS).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z92774; CAB07154.1; -
DR TubercuList: RV3580C; -
DR InterPro: IPR001412; -
DR InterPro: IPR002308; -
DR Pfam: PF01406; TRNA-synt_1e; 1.
DR PRINTS: PR00983; TRNASYNTHCS.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; FALSE NEG.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
FT SITE 35 45 "HIGH" REGION.
FT SITE 267 271 "KMSK" REGION.
FT BINDING 270 270 ATP (BY SIMILARITY).
SO SEQUENCE 469 AA; 51854 MW; 515F9D19482ADC0E CRC64;

Query Match 22.9%; Score 54; DB 1; Length 469;
Best Local Similarity 34.5%; Pred. No. 10;
Matches 10; Conservative 0; Mismatches 7; Indels 12; Gaps 1;

OY 5 PMPWRRKHE-----APEAP 21
DB 89 PWEPMATHERATPAAYDALDVLPSAP 117

RESULT 12
VGL2_CVH22 STANDARD; PRT; 1173 AA.
ID VGL2_CVH22
AC P15423;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUL-1998 (Rel. 38, Last annotation update)
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN).
DE S.
GN Human coronavirus (strain 229E).
OS Human coronavirus (strain 229E).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11137;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90264837; PubMed=2345367;
RA Raabe T., Schelle-Pfanz B., Siddell S.G.;
RT "Nucleotide sequence of the gene encoding the spike glycoprotein of
RT human coronavirus HCV 229E."

```

Query Match	Best Local Similarity	Matches	Score	Pred. No.	DB 1	Length	DB 2	Gaps
4	111	14	22.7%	20.6%	7	990	8	39
176	176	14	22.7%	20.6%	7	990	8	39

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:42:25 ; Search time 48.4 Seconds
(without alignments)
98.409 Million cell updates/sec

Title: SCHN12-444-MODSEQ2.PEP
Perfect score: 236
Sequence: 1 ILRPMWMPMRKHEAPEAPEPIMLKKWMPMRK 36

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_16.*

- 1: sp.archaea.*
- 2: sp.bacteria.*
- 3: sp.fungi.*
- 4: sp.human.*
- 5: sp.invertebrate.*
- 6: sp.mammal.*
- 7: sp.mhc.*
- 8: sp.organelle.*
- 9: sp.phage.*
- 10: sp.plant.*
- 11: sp.potent.*
- 12: sp.unclassified.*
- 13: sp.vertebrate.*
- 14: sp.virus.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73.5	31.1	1245	3 O9Y7V5	O9Y7V5 trichoderma
2	68.5	29.0	746	1 O9JH31	O9JH31 tt virus. o
3	67	28.4	723	14 O9DUC4	O9DUC4 tt virus. o
4	64	27.1	49	14 O9DPM0	O9DPM0 tt virus. o
5	64	27.1	748	14 O9D81	O9D81 tt virus. o
6	63	26.7	147	11 O61427	O61427 mus musculu
7	63	26.7	192	2 O928B7	O928B7 chlamydia p
8	62	26.3	735	14 O9DUC9	O9DUC9 tt virus. o
9	61.5	26.1	2292	14 O66765	O66765 encephalomy
10	61	25.8	252	14 O9IU04	O9IU04 human immun
11	61	25.8	252	14 O9IT26	O9IT26 human immun
12	61	25.8	252	14 O9ITX6	O9ITX6 human immun
13	61	25.8	252	14 O9ITU7	O9ITU7 human immun
14	59.5	25.2	114	2 O9X8C2	O9X8C2 streptomyce
15	59	25.0	95	10 O9ILONO	O9ILONO arabidopsis
16	59	25.0	252	14 O9ITU9	O9ITU9 human immun
17	59	25.0	451	2 O05074	O05074 streptomyce
18	59	25.0	485	2 P72844	P72844 synechocyst
19	58.5	24.8	985	14 O98414	O98414 ovine lentl

20	58	24.6	252	14 O9IU37	O9IU37 human immun
21	58	24.6	252	14 O9IU06	O9IU06 human immun
22	57.5	24.4	376	3 O94516	O94516 schizosacch
23	57	24.2	252	14 O9ITU5	O9ITU5 human immun
24	57	24.2	754	14 O9JH33	O9JH33 tt virus. o
25	57	24.2	970	11 O88821	O88821 mus musculu
26	57	24.2	971	11 O70458	O70458 mus musculu
27	57	24.2	1018	1 O9HXK3	O9HXK3 thermoplas
28	56.5	23.9	157	5 O9Y0E8	O9Y0E8 drosophila
29	56.5	23.9	162	5 O9W1W7	O9W1W7 drosophila
30	56	23.7	273	4 O9N1W2	O9N1W2 homo sapien
31	56	23.7	304	14 O9WCF5	O9WCF5 human immun
32	55	23.5	767	14 O9QUD8	O9QUD8 tt virus. h
33	55.5	23.5	135	14 O9E830	O9E830 anas chloro
34	55.5	23.5	605	14 O9Q5L3	O9Q5L3 vlna virus
35	55.5	23.5	745	14 O9WSV7	O9WSV7 cercophilic
36	55.5	23.5	745	14 O9JG78	O9JG78 tt virus. d
37	55.5	23.5	746	14 O9JG80	O9JG80 tt virus. o
38	55.5	23.5	766	14 O9IFV0	O9IFV0 tt virus. p
39	55.5	23.5	252	14 O9ITV5	O9ITV5 human immun
40	55	23.3	273	4 O99776	O99776 homo sapien
41	55	23.3	352	2 P73267	P73267 synechocyst
42	55	23.3	367	11 O63778	O63778 rattus norv
43	55	23.3	404	14 O9DPJ4	O9DPJ4 human immun
44	55	23.3	404	14 O9DPJ3	O9DPJ3 human immun
45	55	23.3	404	14 O9DPJ3	O9DPJ3 human immun

ALIGNMENTS

RESULT	ID	Query Match	Score	DB 3	Length	1245
O9Y7V5	O9Y7V5	31.1%	73.5	DB 3	Length	1245
AC	O9Y7V5	Best Local Similarity	32.3%	Pred. NO. 0.45		
DT	O1-NOV-1999 (TREMBLrel. 12, Created)	Matches	10	Conservative	1	Mismatches
DT	O1-NOV-1999 (TREMBLrel. 12, Last sequence update)				9	Indels
DE	O1-NOV-1999 (TREMBLrel. 12, Last annotation update)				11	Gaps
DE	CONIDIOSPORE SURFACE PROTEIN.					
GN	CMP1.					
OS	Trichoderma harzianum.					
OC	Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.					
OX	NCBI_TaxID=5544;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	SPPAIN-ATCC 32173;					
RA	Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,					
RA	van Montagu M., Herrera Estrella A., Horwitz B.A.;					
RT	"Developmental regulation of a gene encoding a multidomain					
RT	condiiospore surface protein of Trichoderma, cmp1.";					
RL	Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.					
DR	EMBL: AJ133651; CAB40845.1;					
SO	SEQUENCE 1245 AA; 135824 MW; 3249C749AFRA0CDF8 CRC64;					

Query Match	Score	DB 3	Length	1245
Best Local Similarity	32.3%	Pred. NO. 0.45		
Matches	10	Conservative	1	Mismatches
			9	Indels
			11	Gaps
OY	3 RMPWMPMRKHEAPEAPEPIMLKKWMPMRK 33			
Db	1185 RMPWMPMRKHEAPEAPEPIMLKKWMPMRK 1204			
	-----CWMWMSW 1204			
RESULT	2			
ID	O9JH31	PRELIMINARY:	PRT:	746 AA.
AC	O9JH31			
DT	O1-OCT-2000 (TREMBLrel. 15, Created)			
DT	O1-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DE	O1-OCT-2000 (TREMBLrel. 15, Last annotation update)			
DE	ORF1.			
OS	TT virus.			

RESULT	4	
Q9DT80		
ID	Q9DT80	PRM; 49 AA.
AC	Q9DT80;	
DT	01-MAR-2001 (TREMBLrel. 16, Created)	
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)	

RESULT	6	
061427		
ID	061427	
AC	061427	PRELIMINARY
DT	01-NOV-1996	PRT: 147 AA.
DT	01-NOV-1996	(TREMBlrel. 01, Created)
DT	01-MAR-2001	(TREMBlrel. 01, Last sequence update)
DE	PROCOLLAGEN, TYPE I, ALPHA 1 (ALPHA 1 TYPE I COLLAGEN) (FRAGMENT).	
GN	COL1A1 OR COL1A1.	
OS	Mus musculus (Mouse).	
OC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus	

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=LIVER;
 RA MEDLINE=94344105; PubMed=8065328;
 RX Rhodes K., Rippe R.A., Umezawa A., Nehls M., Brenner D.A., Breindel M.,
 RT "DNA methylation represses the murine alpha 1(I) collagen promoter by
 an indirect mechanism."
 RL Mol. Cell. Biol. 14:5950-5960(1994).
 DR EMBL: X54876; CAA38657.1;
 DR MGD: MGI:88467; COLA1.
 DR InterPro: IPR001007;
 DR Pfam: PF00093; VWC; 1.
 DR PROSITE: PS01208; VMEC; 1.
 DR SMART: SM00214; VMC; 1.
 FT NON_TER 147
 SO SEQUENCE 147 AA; 16652 MW; 9263BFOA91B4307D CRC64;

Query Match 26.7%; Score 63; DB 11; Length 147;
 Best Local Similarity 36.6%; Pred. NO. 1.1;
 Matches 15; Conservative 0; Mismatches 8; Indels 18; Gaps 4;

OY 3 RMPWMPRRKHAEPEAPIMILKKM-----PW-----WPMR 34
 Db 116 RMP--PMTRMH-----PMTAMTSSWSWSPMAPRTPMPMR 147

RESULT 7
 O928B7 PRELIMINARY; PRT: 192 AA.
 AC O928B7;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE CT277 SIMILARITY (CPJ0426 PROTEIN).
 GN CPN0426 OR CPJ0426 OR CPJ0426
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CML029;
 RA MEDLINE=9206606; PubMed=10192388;
 RX Kaiman S., Mitchell W., Marathe R., Lamme C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis."
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shiba T., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kuwara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CML029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Umayam L.A., Ullrich T.,
 RA Berry K., Bass S., Liner K., Weidman J., Khouri H., Craven B.,
 RA Bowman C., Dodson R., Gwin M., Nelson W., Debey R., Kolonay J.,
 RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 pneumoniae AR39."
 RL Nucleic Acids Res. 28:1397-1406(2000).
 DR EMBL: AE001625; AAD18570.1;
 DR EMBL: AP002546; BAA98634.1;
 DR EMBL: AE002194; AAF38182.1;
 DR TIGR: CP0327;

SO SEQUENCE 192 AA; 21464 MW; 9C38C329AEDDB76F CRC64;

Query Match 26.7%; Score 63; DB 2; Length 192;
 Best Local Similarity 42.3%; Pred. NO. 1.4;
 Matches 11; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 4 WPMWMPRRKHAEPEAPIMILKKM 29
 Db 138 WPMWMPRRKHAEPEAPIMILKKM 29

RESULT 8
 O9DUC9 PRELIMINARY; PRT: 735 AA.
 ID O9DUC9;
 AC O9DUC9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE ORF1.
 OS TP virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PT-TTV6;
 RA Okamoto H.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PT-TTV6;
 RX PubMed=11080484;
 RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
 RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
 RT "Species-specific TP viruses in humans and nonhuman primates and their
 phylogenetic relatedness."
 RL Virology 277:368-378(2000).
 DR EMBL: AB041957; BAB19308.1;
 SO SEQUENCE 735 AA; 86132 MW; 9ED818D6BE6FA5D3 CRC64;

Query Match 26.3%; Score 62; DB 14; Length 735;
 Best Local Similarity 27.9%; Pred. NO. 7.2;
 Matches 12; Conservative 4; Mismatches 7; Indels 20; Gaps 3;

OY 2 LRMPWMPRRKHAEPEAPIMILKKM 36
 Db 1 MAMPWMPRRKHAEPEAPIMILKKM 31

RESULT 9
 O66765 PRELIMINARY; PRT: 2292 AA.
 ID O66765;
 AC O66765;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE POLYPROTEIN, COMPLETE CDS.
 OS Enterovirus; Picornaviridae.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirus.
 OX NCBI_TaxID=12104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92148946; PubMed=1310768;
 RX Duke G.M., Hoffman M.A., Palmenberg A.C.;
 RT "Sequence and structural elements that contribute to efficient
 RT encephalomyocarditis virus RNA translation."
 RL J. Virol. 66:1602-1609(1992).
 DR EMBL: M81861; AAA43037.1;
 DR HSSP: P12296; IMEC.
 DR InterPro: IPR000605;
 DR InterPro: IPR001205;

DR InterPro; IPR001676; .
 DR Pfam; PF00073; tnv; 3.
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 KM Polyprotein...
 FT CHAIN 68 137 PROTEIN 1A.
 FT CHAIN 138 393 PROTEIN 1B.
 FT CHAIN 394 624 PROTEIN 1C.
 FT CHAIN 625 901 PROTEIN 1D.
 FT CHAIN 902 1044 PROTEIN 1E.
 FT CHAIN 1045 1194 PROTEIN 2A.
 FT CHAIN 1195 1519 PROTEIN 2B.
 FT CHAIN 1520 1627 PROTEIN 2C.
 FT CHAIN 1628 1832 PROTEIN 3AB.
 FT CHAIN 1833 2292 PROTEIN 3C.
 FT CHAIN 2292 25457 MW; 01C053788CEFC94 CRC64;
 SQ SEQUENCE

Query Match 26.1%; Score 61.5; DB 14; Length 2292;
 Best Local Similarity 27.3%; Pred. No. 26;
 Matches 12; Conservative 6; Mismatches 9; Indels 17; Gaps 1;

QY 3 RRPMPMPRRKHEAP-----MILKKP 29
 DB 967 RRPMPMPRRKHEAP-----MILKKP 1010

RESULT 10
 ID 091U04 PRELIMINARY; PRT; 252 AA.
 AC 091U04
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE V-1 REVERSE TRANSCRIPTASE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID-11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-97001;
 RA Servais J., Lambert C., Fontaine E., Schmit J.C.;
 RT "Gradual emergence and clinical relevance of HIV protease inhibitors
 resistance in HAART failure patients."
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ401761; CAB86530.1;
 DR InterPro: IPR000477;
 DR Pfam; PF00078; rvt; 1.
 KM RNA-directed DNA polymerase.
 FT NON_TER 1
 FT NON_TER 252
 SQ SEQUENCE 252 AA; 29441 MW; F2B8FB9A8334D9C CRC64;

Query Match 25.8%; Score 61; DB 14; Length 252;
 Best Local Similarity 27.3%; Pred. No. 3.3;
 Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;

QY 1 IIRPMPMPRRKHEAP-----EAPIMILKKP 30
 DB 209 IIRPMPMPRRKHEAP-----EAPIMILKKP 252

RESULT 11
 ID 091T26 PRELIMINARY; PRT; 252 AA.
 AC 091T26
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE V-1 REVERSE TRANSCRIPTASE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.

OX NCBI_TaxID-11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-97095;
 RA Servais J., Lambert C., Fontaine E., Schmit J.C.;
 RT "Gradual emergence and clinical relevance of HIV protease inhibitors
 resistance in HAART failure patients."
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ401769; CAB86538.1;
 DR InterPro: IPR000477;
 DR Pfam; PF00078; rvt; 1.
 KM RNA-directed DNA polymerase.
 FT NON_TER 1
 FT NON_TER 252
 SQ SEQUENCE 252 AA; 29346 MW; 7CB1C1FCD16750CF CRC64;

Query Match 25.8%; Score 61; DB 14; Length 252;
 Best Local Similarity 27.3%; Pred. No. 3.3;
 Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;

QY 1 IIRPMPMPRRKHEAP-----EAPIMILKKP 30
 DB 209 IIRPMPMPRRKHEAP-----EAPIMILKKP 252

RESULT 12
 ID 091T26 PRELIMINARY; PRT; 252 AA.
 AC 091T26
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE V-1 REVERSE TRANSCRIPTASE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID-11676;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-97451;
 RA Servais J., Lambert C., Fontaine E., Schmit J.C.;
 RT "Gradual emergence and clinical relevance of HIV protease inhibitors
 resistance in HAART failure patients."
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AJ401789; CAB86558.1;
 DR InterPro: IPR000477;
 DR Pfam; PF00078; rvt; 1.
 KM RNA-directed DNA polymerase.
 FT NON_TER 1
 FT NON_TER 252
 SQ SEQUENCE 252 AA; 29450 MW; 8E48D6280726D12 CRC64;

Query Match 25.8%; Score 61; DB 14; Length 252;
 Best Local Similarity 27.3%; Pred. No. 3.3;
 Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;

QY 1 IIRPMPMPRRKHEAP-----EAPIMILKKP 30
 DB 209 IIRPMPMPRRKHEAP-----EAPIMILKKP 252

RESULT 13
 ID 091T27 PRELIMINARY; PRT; 252 AA.
 AC 091T27
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
 DE V-1 REVERSE TRANSCRIPTASE (FRAGMENT).
 OS Human immunodeficiency virus type 1.
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
 OX NCBI_TaxID-11676;

RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=98324;
 RC "Gradual emergence and clinical relevance of HIV protease inhibitors
 RA Servais J., Lambert C., Fontaine E., Schmit J.C.;
 RT resistance in HAART failure patients";
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ401818; CAB86587.1;
 DR InterPro: IPR000477;
 DR Pfam: PF00078; rvc; 1;
 DR RNA-directed DNA polymerase.
 KM NON_TER 1
 FT NON_TER 1
 SO SEQUENCE 252 AA; 29495 MW; 33AD2C43DE2E436 CRC64;

Query Match 25.8%; Score 61; DB 14; Length 252;
 Best Local Similarity 27.3%; Pred. No. 3.3; 10; Indels 14; Gaps 1;
 Matches 12; Conservative 8; Mismatches 10; Indels 14; Gaps 1;
 OY 1 ILRPMWPMRRKHEAP-----EAPIMILKMPW 30
 DB 209 LIRGVTTPDKKHQKPPFLIMGYELHPDKVTQPIYLPEKDSW 252

RESULT 14
 ID 09X8C2 PRELIMINARY; PRT: 114 AA.
 AC 09X8C2:
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE HYPOTHETICAL 13.0 KDA PROTEIN.
 GN SCE36.09.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Oliver K., Harris D.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RL the 8 Mb Streptomyces coelicolor A3(2) chromosome";
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Bentley S.D.; Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RC MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RL the 8 Mb Streptomyces coelicolor A3(2) chromosome";
 RL MOL. Microbiol. 21:77-96(1996).
 DR EMBL; ADO49763; CAB42078.1;
 DR Hypothetical protein.
 KM SEQUENCE 114 AA; 13031 MW; B28E223FC4A0BB49 CRC64;

Query Match 25.2%; Score 59.5; DB 2; Length 114;
 Best Local Similarity 50.0%; Pred. No. 2.3; 7; Indels 1; Gaps 1;
 Matches 10; Conservative 2; Mismatches 10; Indels 1; Gaps 1;

OY 17 PEAPIMILKMW-PWMPWR 35
 DB 93 PETAFADARRWRPRMPWR 112

RESULT 15
 ID 09LONO PRELIMINARY; PRT: 95 AA.
 AC 09LONO:
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE F5D14.5.
 GN F5D14.5.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 OC Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Liu S.X., Chan A., Yu G., Lee J.M., Lenz C., Pham P., Sakano H.,
 RA Toriumi M., Vysotskaya V.S., Chin C., Chlou J., Choi E., Chung M.,
 RA Gonzalez A., Howng B., Liu A., Vayenberg M., Altafi H., Brooks S.,
 RA Buehler E., Chao Q., Conn L., Conway A.B., Hansen N.F.,
 RA Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M., Nguyen M.,
 RA Palm C.J., Shinn P., Southwick A., Davis R.W., Ecker J.R.,
 RA Federpsiel N.A., Theologis A.;
 RT "The sequence of BAC F5D14 from Arabidopsis thaliana chromosome 1.";
 RL Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (Jun-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Theologis A.;
 RL Submitted (Jul-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007767; AAF81325.1;
 DR SEQUENCE 95 AA; 10624 MW; 0C3985771E8B54E CRC64;

Query Match 25.0%; Score 59; DB 10; Length 95;
 Best Local Similarity 21.3%; Pred. No. 2.2; 4; Indels 28; Gaps 3;
 Matches 10; Conservative 5; Mismatches 10; Indels 28; Gaps 3;

OY 4 WP-----WMPRRKHEAPPAEPIMILK-----KWPMPW 33
 DB 46 WPVYVVAVGCGGRMMW-----PVLITDVGEMSMW 81

Search completed: June 21, 2001, 08:42:26
 Job time: 277 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 21, 2001, 08:34:03 ; Search time 42.99 Seconds

(without alignments)
49.357 Million cell updates/sec

Title: SCHIN2-444-MODSEQ2B.PEP
Perfect score: 231

Sequence: 1 ILRWPMPWRKHEAPEAPEIPIMLRPMPWRK

Scoring table: BL0SUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

1: /SID8/gcgdata/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162	70.1	27	19	AAW6363
2	154	66.7	28	21	AAV91800
3	128	55.4	63	21	AAV44668
4	128	55.4	63	21	AAV57142
5	95	41.1	15	19	AAW6360
6	95	41.1	15	21	AAV91784
7	90.5	39.2	21	19	AAV24552
8	90.5	39.2	21	19	AAW6376
9	90.5	39.2	21	19	AAV91796
10	86	37.2	12	21	AAV24550
11	86	37.2	12	21	AAV94496

12	86	37.2	12	21	AAV91791	Amino acid sequenc
13	86	37.2	16	18	AAW12882	Antimicrobial cati
14	86	37.2	20	19	AAV24553	Indolicidin analog
15	86	37.2	20	21	AAV91797	Amino acid sequenc
16	86	37.2	21	19	AAV24554	Indolicidin analog
17	86	37.2	21	21	AAV91798	Amino acid sequenc
18	84	36.4	16	18	AAW12899	Antimicrobial cati
19	83	35.9	12	19	AAV24567	Indolicidin analog
20	83	35.9	12	21	AAV91788	Amino acid sequenc
21	82	35.5	12	19	AAV24594	Indolicidin analog
22	82	35.5	12	19	AAW6364	Indolicidin analog
23	82	35.5	12	21	AAV91817	Amino acid sequenc
24	82	35.5	12	21	AAV91841	Amino acid sequenc
25	81	35.1	12	19	AAV24605	Indolicidin analog
26	81	35.1	12	19	AAV24595	Indolicidin analog
27	81	35.1	12	19	AAW6361	Indolicidin analog
28	81	35.1	12	21	AAV91785	Amino acid sequenc
29	81	35.1	12	21	AAV91842	Amino acid sequenc
30	81	35.1	12	21	AAV91852	Amino acid sequenc
31	80	34.6	12	21	AAV24596	Indolicidin analog
32	80	34.6	12	19	AAV24603	Indolicidin analog
33	80	34.6	12	19	AAV24604	Indolicidin analog
34	80	34.6	12	21	AAV91843	Amino acid sequenc
35	80	34.6	12	21	AAV91850	Amino acid sequenc
36	80	34.6	12	21	AAV91851	Amino acid sequenc
37	79	34.2	13	19	AAV24565	Indolicidin analog
38	79	34.2	13	21	AAV91786	Amino acid sequenc
39	78	33.8	12	19	AAV24598	Indolicidin analog
40	78	33.8	12	19	AAV24601	Indolicidin analog
41	78	33.8	12	21	AAV91845	Amino acid sequenc
42	78	33.8	12	21	AAV91848	Amino acid sequenc
43	77.5	33.5	13	18	AAW12896	Antimicrobial cati
44	77.5	33.5	13	19	AAV24613	Indolicidin analog
45	77.5	33.5	13	21	AAV91803	Amino acid sequenc

ALIGNMENTS

RESULT 1	
ID	AAW6363
AAW6363	standard; peptide; 27 AA.
XX	
AC	AAW6363;
XX	
DT	12-JAN-1999 (first entry)
XX	
DE	Indolicidin analogue MBI 11B20.
XX	
KW	Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW	bacterial infection; tolerance; antibacterial; microorganism;
XX	bacteria; fungus; parasite; virus.
XX	
OS	Bos taurus.
OS	Synthetic.
XX	
PN	WO9840401-A2.
XX	
PD	17-SEP-1998.
XX	
PF	10-MAR-1998; 98MO-CA00190.
XX	
PR	25-FEB-1998; 98US-0030619.
PR	10-MAR-1997; 97US-0040649.
PR	20-AUG-1997; 97US-0915314.
XX	26-SEP-1997; 97US-0060099.
XX	
PA	(MICR-) MICROLOGIX BIOTECH INC.
XX	
PI	Fraser JR, McNicol PJ, West MHP;
XX	
DR	WPI; 1998-520800/44.
XX	

PT New indolicidin peptide analogues - useful for, e.g. enhancing
 PT activity of antibiotic or overcoming tolerance, acquired resistance
 or inherent resistance of microorganisms

PS Claim 1; Page 91; 105pp; English.

XX The present sequence represents an indolicidin analogue. The present
 CC invention describes compositions and methods for treating infection,
 CC especially bacterial infections. The compositions and methods use
 CC cationic peptides in combination with an antibiotic agent which are
 CC then administered to a patient to enhance the activity of the antibiotic
 CC agent. to overcome: (a) tolerance; (b) acquired resistance; and (c)
 CC inherent resistance. The combinations of antibiotics and cationic
 CC peptides can provide synergistic activity against a microorganism that
 CC is tolerant, inherently resistant, or has acquired resistance to an
 CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 CC parasites and viruses.

XX Sequence 27 AA:

Query Match 70.1%; Score 163; DB 19; Length 27;
 Best Local Similarity 71.4%; Pred. No. 3.7e-14;
 Matches 25; Conservative 0; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILRPMWPMRKRHEAPEAPIMILRPMWPMRKR 35
 1 ILRPMWPMRKRHEAPEAPIMILRPMWPMRKR 25

RESULT 2

AAV91800 standard; Peptide; 28 AA.

AAV91800;

06-JUN-2000 (first entry)

Amino acid sequence of cationic peptide MBI 11B20CN.

Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 multidrug resistance.

Synthetic.

WO9965506-A2.

23-DEC-1999.

14-JUN-1999; 99WO-CA00552.

12-JUN-1998; 98US-0096541.

(MICR-) MICROLOGIX BIOTECH INC.

Friedland HD, Krieger TJ, Taylor R, Erile D, Fraser JR, West MP;

WPI; 2000-223549/19.

Novel pharmaceutical composition containing optionally activated
 polyoxalkylene-modified cationic peptides, useful for treating tumours

Claim 1; Page 15; 94pp; English.

This sequence represents a cationic peptide amino acid sequence, which
 can be used in the pharmaceutical composition of the invention. The
 invention relates to a pharmaceutical composition containing at least one
 activated polyoxalkylene (APO)-modified cationic peptide. The
 modification of peptides with APO increases their activity against tumour
 cells, including those with a multidrug resistant phenotype. The

CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

XX Sequence 28 AA:

Query Match 66.7%; Score 154; DB 21; Length 28;
 Best Local Similarity 62.9%; Pred. No. 3.9e-13;
 Matches 22; Conservative 3; Mismatches 0; Indels 10; Gaps 1;

OY 1 ILRPMWPMRKRHEAPEAPIMILRPMWPMRKR 35
 1 ILRPMWPMRKRHEAPEAPIMILRPMWPMRKR 25

AAV44668 standard; Protein; 63 AA.

AAV44668;

18-APR-2000 (first entry)

poly-(Indol (1-13)-Met-Ala-Arg-Ile-Ala-Met)3 protein.

Crosslinked indolicidin analog; X-indolicidin; poly-indol 1-13;
 stability; bovine neutrophil; antimicrobial; antibacterial; fungicide;
 KM protozoicide; virucide; anti-HIV; human immunodeficiency virus-1;
 KM HIV-1; gram positive bacteria; gram negative; Staphylococcus aureus;
 KM Escherichia coli; Salmonella typhimurium; yeast; fungi; protozoa;
 KM Candida albicans; Cryptococcus neoformans; Giardia; Acanthamoeba;
 hexapeptide spacer.

Synthetic.

Bos sp.

Location/Qualifiers

1..5 /label- Enterokinase_recognition_site

5..6 /label- Enterokinase_cleavage_site

6..7 /label- Cyanogen_bromide_cleavage_site

20..21 /label- Cyanogen_bromide_cleavage_site

25..26 /label- Cyanogen_bromide_cleavage_site

39..40 /label- Cyanogen_bromide_cleavage_site

44..45 /label- Cyanogen_bromide_cleavage_site

58..59 /label- Cyanogen_bromide_cleavage_site

20..25 /label- Hexapeptide_spacer

39..44 /label- Hexapeptide_spacer

58..63 /label- Hexapeptide_spacer

WO9965510-A1.

23-DEC-1999.

20-MAY-1999; 99WO-US11165.

18-JUN-1998; 98US-0096631.

(REGC) UNIV CALIFORNIA.

Seleted ME, Osapay K;

```

DR  WPI: 2000-147133/13.
DR  N-PSDB: AA249764.
XX
PT  Crosslinked indolicidin analogs with antimicrobial activity against
XX  bacteria, yeast, fungi, protozoa and viruses
XX
PS  Example 1C: Fig 1, 53pp: English.
XX
CC  The patent discloses crosslinked analogs of indolicidin (Indol 1-13)
CC  which is a naturally occurring peptide isolated from bovine neutrophils
CC  and has antimicrobial activity. The crosslinked indolicidin
CC  (X-indolicidin) analogs are stable and have antimicrobial activity
CC  against gram positive and negative bacteria (e.g. Staphylococcus aureus,
CC  Escherichia coli and Salmonella typhimurium), yeasts and fungi (e.g.
CC  Candida albicans, Cryptococcus neoformans), protozoa (e.g. Giardia
CC  species and Acanthamoeba species), and viruses (e.g. HIV-1).
CC  They can be used for reducing or inhibiting the growth or survival of
CC  microorganisms in an environment e.g. a food or food product, a
CC  solution, an inanimate object comprising a surface, or a mammal.
CC  The present sequence is a protein comprising three
CC  copies of Indol 1-13 each separated by a hexapeptide spacer sequence.
CC  A recombinant construct encoding this sequence was used for the
CC  expression of Indol-homoserine (Hse) analog. The ability of
CC  Indol-Hse analog to maintain antimicrobial activity provides a means to
CC  produce X-indolicidin analog precursors in sufficient quantities.
XX
SQ  Sequence 63 AA:
XX
Query Match 55.4%; Score 128; DB 21; Length 63;
Best Local Similarity 58.8%; Pred. No. 1.8e-09;
Matches 20; Conservative 2; Mismatches 4; Indels 8; Gaps 2;
Oy 3 RWPMPWRRKHAEPEAEPIML--RWPMPWRR 34
:||||||| | |||:|||||||
Db 11 kwpmpwrrm-----ariamlpwkwpmpwrr 38
XX
RESULT 4
AA57142
ID AA57142 standard; Protein: 63 AA.
XX
AC AA57142;
XX
DT 28-FEB-2000 (first entry)
XX
DE Indolicidin fusion peptide amino acid sequence.
XX
KW Indolicidin analogue; antimicrobial activity; helminth; bacteria; virus;
KW treatment; inhibit growth; micro-organism; contact lens solution;
KW transgenic plant; surgical instrument; yeast; fungi; protozoa.
XX
OS Synthetic.
XX
PN WO9558141-A1.
XX
PD 18-NOV-1999.
XX
PF 05-MAY-1999; 99WO-US09942.
XX
PR 12-MAY-1998; 98US-0076227.
XX
PA (REGC ) UNIV CALIFORNIA.
XX
PI Selected ME:
XX
XX WPI: 2000-053028/04.
DR N-PSDB: AA245123.
XX
PT New indolicidin analogues, active against bacteria, yeast, fungi,
PT protozoa and virus, used for, e.g. treating infections -
PS Disclosure: Fig 6; 62pp: English.

```

```

XX
CC  This is the amino acid sequence of an example of a fusion protein which
CC  consists of an indolicidin analogue linked to another peptide.
CC  Peptides AA57109-Y57138 and AA57143-Y57144 are new indolicidin
CC  analogues, which have a homoserine residue and/or a truncated amino
CC  terminal region. The analogues have the following amino acid sequence:
CC  Xaa1-Xaa2-Xaa3-Xaa4-Xaa5-Xaa6-Pro-Xaa6-Xaa6-Pro-Xaa6-Xaa7-Xaa7-Xaa8
CC  where:
CC  Xaa1 = Ile, Leu, Val, Ala, Gly or absent;
CC  Xaa2 = Ile, Leu, Val, Ala, Gly or absent;
CC  Xaa3 = Pro or absent;
CC  Xaa4 = Trp, Phe or absent;
CC  Xaa5 = Arg, Lys or absent;
CC  Xaa6 = Trp or Phe;
CC  Xaa7 = Arg, Lys or absent;
CC  Xaa8 = homoserine (Hse), Met, Met-Xaa9-Met or absent, and
CC  Xaa9 = at least one amino acid;
CC  provided that if Xaa1 is present, Xaa8 = Hse, Met or Met-Xaa9-Met;
CC  and further provided that: if Xaa2 is absent, Xaa1 is absent; if Xaa3 is
CC  absent, Xaa1 and Xaa2 are absent; if Xaa4 is absent, Xaa1, Xaa2 and Xaa3
CC  are absent; and if Xaa5 is absent, Xaa1, Xaa2, Xaa3 and Xaa4 are absent.
CC  The indolicidin analogues can be used to create a fusion polypeptide
CC  consisting of the analogue linked to a peptide. The indolicidin
CC  analogues have antimicrobial activity against gram positive bacteria,
CC  gram negative bacteria, yeast, fungus, protozoa and viruses (e.g. HIV-1).
CC  They are also active against helminths. The analogues can be used for
CC  reducing or inhibiting growth or survival of a microorganism. They can be
CC  used for treating infections. They can also be included in a liquid such
CC  as water or an aqueous solution, e.g. contact lens solution. The
CC  analogues have potential uses in food products, and in objects such as
CC  the surface of an instrument used to prepare food or to perform surgery.
CC  Transgenic plants or animals useful in the food industry can be produced
CC  by introducing a nucleic acid molecule encoding an indolicidin analogue
CC  into the germline cells of such organisms.
XX
SQ  Sequence 63 AA:
XX
Query Match 55.4%; Score 128; DB 21; Length 63;
Best Local Similarity 58.8%; Pred. No. 1.8e-09;
Matches 20; Conservative 2; Mismatches 4; Indels 8; Gaps 2;
Oy 3 RWPMPWRRKHAEPEAEPIML--RWPMPWRR 34
:||||||| | |||:|||||||
Db 11 kwpmpwrrm-----ariamlpwkwpmpwrr 38
XX
RESULT 5
AAW6360
ID AAW6360 standard; peptide: 15 AA.
XX
AC AAW6360;
XX
DT 12-JAN-1999 (first entry)
XX
DE Indolicidin analogue MBI 11A9.
XX
KW Indolicidin analogue; resistance; cationic peptide; antiprotic;
KW bacterial infection; tolerance; antibacterial; microorganism;
KW bacteria; fungus; parasite; virus.
XX
OS Synthetic.
XX
PN WO9840401-A2.
XX
PD 17-SEP-1998.
XX
PF 10-MAR-1998; 98WO-CA00190.
XX
PR 25-FEB-1998; 98US-0030619.
PR 10-MAR-1997; 97US-0040649.
PR 20-AUG-1997; 97US-0915314.

```


CC controlled are leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
CC derived from the analogues may be used similarly; the compounds may
CC also be prepared from antibiotics or antiarrhythmic agents. The analogues
CC may be used therapeutically or to coat medical devices; also they are
CC useful as surface disinfectants, as additives to shampoo or soaps, as
CC insecticides or herbicides, or as preservatives for foods and technical
CC materials. The analogues are administered by injection, lavage, orally
CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
CC spectrum of activity than indolicidin and modification as compounds
CC reduces their toxicity.

XX
SQ Sequence 21 AA:

Query Match 39.2%; Score 90.5; DB 19; Length 21;
Best Local Similarity 61.5%; Pred. No. 2.8e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

Oy 1 ILRPPMPWRRKHEAPEAPIMILR 26
Db 1 ILRPPMPWRRKHEAPEAPIMILR 17

RESULT 8

AAW6376
ID AAW6376 standard; peptide: 21 AA.

AC AAW6376;

XX 12-JAN-1999 (first entry)

DE Cationic peptide of claim 15 #3.

XX Indolicidin analogue; resistance; cationic peptide; antibiotic;

KW bacterial infection; tolerance; antibacterial; microorganism;

KN bacteria; fungus; parasite; virus.

OS Synthetic.

PN MO9840401-A2.

PD 17-SEP-1998.

XX 10-MAR-1998; 98WO-CA00190.

PR 25-FEB-1998; 98US-0030619.

PR 10-MAR-1997; 97US-0040649.

PR 20-AUG-1997; 97US-0915314.

PR 26-SEP-1997; 97US-0060099.

XX (MICR-) MICROLOGIX BIOTECH INC.

PI Fraser JR, McNicol PJ, West MHP;

DR WPI: 1998-520800/44.

XX New indolicidin peptide analogues - useful for, e.g. enhancing

PT activity of antibiotic or overcoming tolerance; acquired resistance

PT or inherent resistance of microorganisms

XX Claim 15; Page 93; 105pp; English.

CC The present sequence represents a specifically claimed cationic peptide
CC from the present invention. The present invention describes compositions
CC and methods for treating infection, especially bacterial infections. The
CC compositions and methods use cationic peptides in combination with an
CC antibiotic agent which are then administered to a patient to enhance the
CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
CC acquired resistance; and (c) inherent resistance. The combinations of
CC antibiotics and cationic peptides can provide synergistic activity
CC against a microorganism that is tolerant, inherently resistant, or has

CC acquired resistance to an antibiotic agent. They can be used for killing
CC e.g. bacteria, fungi, parasites and viruses.

XX
SQ Sequence 21 AA:

Query Match 39.2%; Score 90.5; DB 19; Length 21;
Best Local Similarity 61.5%; Pred. No. 2.8e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

Oy 1 ILRPPMPWRRKHEAPEAPIMILR 26
Db 1 ILRPPMPWRRKHEAPEAPIMILR 17

RESULT 9

AA91796
ID AA91796 standard; Peptide: 21 AA.

AC AA91796;

XX 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11B16CN.

XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;

KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;

KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

XX multidrug resistance.

OS Synthetic.

PN WO9965506-A2.

PD 23-DEC-1999.

XX 14-JUN-1999; 99WO-CA00552.

XX 12-JUN-1998; 98US-0096541.

XX (MICR-) MICROLOGIX BIOTECH INC.

PI Friedland HD, Krieger TJ, Taylor R, Erife D, Fraser JR, West MHP;

DR WPI: 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated

PT polyoxalkylene-modified cationic peptides, useful for treating tumours

PT Dislosure; Page 15; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which

CC can be used in the pharmaceutical composition of the invention. The

CC invention relates to a pharmaceutical composition containing at least one

CC activated polyoxalkylene (APO)-modified cationic peptide. The

CC modification of peptides with APO increases their activity against tumour

CC cells, including those with a multidrug resistant phenotype. The

CC pharmaceutical composition can be used to treat tumours, specifically

CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,

CC cervix, uterus, skin, prostate, liver and colon.

XX
SQ Sequence 21 AA:

Query Match 39.2%; Score 90.5; DB 21; Length 21;
Best Local Similarity 61.5%; Pred. No. 2.8e-05;
Matches 16; Conservative 1; Mismatches 0; Indels 9; Gaps 1;

Oy 1 ILRPPMPWRRKHEAPEAPIMILR 26
Db 1 ILRPPMPWRRKHEAPEAPIMILR 17

RESULT 10
 AAY24550
 ID AAY24550 standard; peptide: 12 AA.
 XX
 AC AAY24550;
 XX
 DT 18-AUG-1999 (first entry)
 XX
 DE Indolicidin analogue #2.
 XX
 DE Indolicidin; bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antiaerobic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KW food; technical material.
 XX
 OS Synthetic.
 XX
 PN WO9807745-A2.
 XX
 PD 26-FEB-1998.
 XX
 PF 21-AUG-1997; 97WO-US14779.
 XX
 PR 13-JAN-1997; 97US-0034949.
 XX
 PR 21-AUG-1996; 96US-0024754.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Erfile D, Fraser JR, Krieger TJ, Taylor R, West MH;
 XX
 DR WPI: 1998-169090/15.
 XX
 PT New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 XX
 PS Claim 11; Page 88; 129pp; English.
 XX
 CC AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VIII) containing up to 25 amino acids (aa): RZXZXZXB (I), BXZXZXZXB
 CC (II), BBZXZXZXB (III), BXZXZXZBBn(AA)nMILBAGS (IV), BXZXZXZBB (V),
 CC (V), LBnZXZXZBBn (VI), LKZXZXZBBn (VII) and BXZXZXZBB (VIII).
 CC where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 XX
 SQ Sequence 12 AA:

Query Match 37.2%; Score 86; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5.5e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRPMWMPWRK 12
 DB 1 ILRPMWMPWRK 12

RESULT 11
 AAY94496
 ID AAY94496 standard; Peptide: 12 AA.
 XX
 AC AAY94496;
 XX
 DT 20-SEP-2000 (first entry)
 XX
 DE MBI-11B7 peptide derived from indolicidin.
 XX
 DE Cellulose binding domain; CBD; cationic peptide;
 KW MBI-11B7; indolicidin; bovine.
 XX
 OS Bos taurus.
 XX
 PN WO200031279-A2.
 XX
 PD 02-JUN-2000.
 XX
 PF 19-NOV-1999; 99WO-CA01107.
 XX
 PR 20-NOV-1998; 98US-0109218.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Burian J, Bartfeld D;
 XX
 DR WPI: 2000-400086/34.
 XX
 PT Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products -
 XX
 PS Disclosure; Page 24; 73pp; English.
 XX
 CC A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a promoter and
 CC multi-domain fusion protein expression cassette comprising a promoter and
 CC a nucleic acid molecule expressed as an insoluble protein. The inclusion
 CC of anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the
 CC fusion protein is controlled. This cassette allows high yield, stable
 CC production of the cationic peptide. Cationic peptides such as
 CC bovine indolicidin may be used as antimicrobial agents. The present
 CC sequence is the MBI-11B7 peptide. MBI-11B7 is a cationic peptide derived
 CC from modifications of indolicidin.
 XX
 SQ Sequence 12 AA:

Query Match 37.2%; Score 86; DB 21; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5.5e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRPMWMPWRK 12
 DB 1 ILRPMWMPWRK 12

RESULT 12
 AAY91791
 ID AAY91791 standard; Peptide: 12 AA.
 XX
 AC AAY91791;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Amino acid sequence of cationic peptide MBI 11B7CN.
 XX
 DE Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

KM multidrug resistance.
 XX Synthetic.
 OS
 XX W09965506-A2.
 PN
 XX
 XX 23-DEC-1999.
 PD
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 XX 12-JUN-1998; 98US-0096541.
 PR
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI: 2000-223549/19.
 DR
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PT polyoxalylene-modified cationic peptides, useful for treating tumours
 PT
 PS Claim 1; Page 14; 94pp; English.
 XX
 XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SQ Sequence 12 AA:
 XX
 XX
 Query Match 37.2%; Score 86; DB 21; Length 12;
 Best Local Similarity 100.0%; Pred. No. 5.5e-05;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILRWPMWPMWRK 12
 |||||
 Db 1 ILRWPMWPMWRK 12
 |||||
 RESULT 13
 AAM12882
 ID AAM12882 standard; peptide: 16 AA.
 XX
 AC AAM12882:
 XX
 DT 10-DEC-1997 (first entry)
 XX
 DE Antimicrobial cationic peptide SEQ ID NO:11.
 XX
 KM Bacterial; viral; antitumour; food; preservative; inhibitor; growth;
 KM bacterium; yeast; endotoxaemia; sepsis; antibiotic; fungal;
 KM antiviral; Candida albicans; steriliant; Salmonella; Yersinia;
 KM Shigella.
 XX
 OS Synthetic.
 XX
 PN W09708199-A2.
 XX
 PD 06-MAR-1997.
 XX
 PF 23-AUG-1996; 96WO-IB00996.
 XX
 PR 23-AUG-1995; 95US-0002687.
 XX
 PA (UYBR-) UNIV BRITISH COLUMBIA.
 XX

PI Falla TJ, Gough M, Hancock REM;
 XX
 DR WPI: 1997-179179/16.
 XX
 PT Cationic peptide(s) having anti-microbial activity - used for the
 PT inhibition of bacterial and viral growth, as an antitumour agent,
 PT and as a food preservative
 PT
 PS Claim 2; Page 66; 89pp; English.
 XX
 XX The present sequence represents a specifically claimed novel isolated
 CC cationic peptide which has antimicrobial activity. The amino acid
 CC sequence of antimicrobial cationic peptides (including the present
 CC sequence) is selected from: X1X1ProX3X2Pro(X2X2Pro)nX2X3(X5)O;
 CC X1X1ProX2X3X4(X5)ProX2X3X3; X1X1X3(ProTrp)uX3X2X5X2X5X2(X5)O;
 CC X1X1X3X3X2Pro(X2X2Pro)nX2(X5)m; where m = 1-5; n = 1-2; o = 2-5; r
 CC = 0-8; u = 0-1; X1 = Ile, Leu, Val, Phe, Tyr, Trp or Met; X2 = Trp or
 CC Phe; X3 = Arg or Lys; X4 = Trp or Lys; and X5 = Phe, Trp, Arg, Lys or
 CC Pro. The peptides are preferably amidated or carboxymethylated. The
 CC peptides may be used in methods for inhibiting the growth of a bacterium
 CC or yeast, or for inhibiting an endotoxaemia or sepsis associated
 CC disorder in a subject. The peptides have a broad activity against
 CC antibiotic resistant bacteria, combined with activity against the
 CC medically important fungus Candida albicans. In addition, the peptides
 CC are useful as antitumour agents and/or antiviral agents. The peptides
 CC may be used as sterilants or preservatives of materials susceptible to
 CC microbial or viral contamination, e.g. in processed foods to inhibit
 CC Salmonella, Yersinia and Shigella. The peptides are compact and tend to
 CC have a unique polypeptide type II extended helix structure that permits
 CC them to span the membrane with relatively few amino acids. The peptides
 CC possess the ability to work synergistically with antibiotics, and in
 CC addition, some of them possess anti-endotoxin activity.
 CC
 SQ Sequence 16 AA:
 XX
 XX
 Query Match 37.2%; Score 86; DB 18; Length 16;
 Best Local Similarity 41.9%; Pred. No. 7.6e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 18; Gaps 2;
 QY 5 PM-WPMRKHAEPEAPIMILRWPMWPMWR 34
 |||||
 Db 3 PMWPMWPMWR-----WPMWPMWR 16
 |||||
 RESULT 14
 AAY24553
 ID AAY24553 standard; peptide: 20 AA.
 XX
 AC AAY24553:
 XX
 DT 18-AUG-1999 (first entry)
 XX
 DE Indolicidin analogue #5.
 XX
 KM Indolicidin; bacterial infection; photo-oxidised solubiliser;
 KM antimicrobial; antibiotic; antidiarrhythmic; surface disinfectant;
 KM additive; shampoo; soap; insecticide; herbicide; preservative;
 KM food; technical material.
 XX
 OS Synthetic.
 XX
 PN W09807745-A2.
 XX
 PD 26-FEB-1998.
 XX
 PF 21-AUG-1997; 97WO-US14779.
 XX
 PR 13-JAN-1997; 97US-0034949.
 XX
 PR 21-AUG-1996; 96US-0024754.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX

Claim 11; Page 88; 129pp; English.

Sequence 20 AA;

Matches 12; Conservation

1 11rwprwprk 12

RESULT 15

ID	AA	Y91797	standard;	Peptide;	20	AA

AC AAY91797

DT 06-JUN-2000 (first entry)

AA Amino acid sequence of cationic peptide MBI 11B17CN.
DE

multidrug resistance.

OS Synthetic

PN W09965506-A2.

PD 23-DEC-1999

PF 14-JUN-1999; 99WO-CA005552.

PR 12-JUN-1998; 98US-0096541

PA (MICR-) MICROLOGIX BIOTECH INC.

XX

DR WPI; 2000-223549/19.

PS Disclosure; Page 15; 94pp; English.

50	Sequence	20 AA
----	----------	-------

Matches 12; Conserva

Db 1 111rwprwprk 12

Job time: 266 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:35:23 ; Search time 25.06 Seconds
(Without Alignments)
28.135 Million cell updates/sec

Title: SCHNIZ-444-MODSEQ2B.PEP
Perfect score: 231
Sequence: 1 ILRMPMPWRRKHEAPEAPIMILRMPMPWRRK 35

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 2014635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/prodata/2/1aa/5A-COMB.pep:*
- 2: /cgn2_6/prodata/2/1aa/5B-COMB.pep:*
- 3: /cgn2_6/prodata/2/1aa/6A-COMB.pep:*
- 4: /cgn2_6/prodata/2/1aa/6B-COMB.pep:*
- 5: /cgn2_6/prodata/2/1aa/PCMTUS-COMB.pep:*
- 6: /cgn2_6/prodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.5	39.2	21	4	US-08-915-314-46 Sequence 46, Appl
2	86	37.2	12	4	US-08-915-314-42 Sequence 42, Appl
3	86	37.2	16	4	US-08-702-054B-11 Sequence 11, Appl
4	86	37.2	20	4	US-08-915-314-47 Sequence 47, Appl
5	86	37.2	21	4	US-08-915-314-48 Sequence 48, Appl
6	84	36.4	16	4	US-08-702-054B-38 Sequence 38, Appl
7	83	35.9	12	4	US-08-915-314-40 Sequence 40, Appl
8	83	35.5	12	4	US-08-915-314-76 Sequence 76, Appl
9	81	35.1	12	4	US-08-915-314-77 Sequence 77, Appl
10	81	35.1	12	4	US-08-915-314-87 Sequence 87, Appl
11	80	34.6	12	4	US-08-915-314-78 Sequence 78, Appl
12	80	34.6	12	4	US-08-915-314-85 Sequence 85, Appl
13	80	34.6	12	4	US-08-915-314-86 Sequence 86, Appl
14	79	34.2	13	4	US-08-915-314-38 Sequence 38, Appl
15	78	33.8	12	4	US-08-915-314-83 Sequence 83, Appl
16	78	33.8	12	4	US-08-915-314-80 Sequence 80, Appl
17	77.5	33.5	13	4	US-08-915-314-51 Sequence 51, Appl
18	77.5	33.5	13	4	US-08-702-054B-34 Sequence 34, Appl
19	77	33.3	12	4	US-08-915-314-69 Sequence 69, Appl
20	76	32.9	13	4	US-08-915-314-25 Sequence 25, Appl
21	76	32.9	13	4	US-08-915-314-30 Sequence 30, Appl
22	76	32.9	13	4	US-08-915-314-62 Sequence 62, Appl
23	76	32.9	13	4	US-08-915-314-63 Sequence 63, Appl
24	76	32.9	13	4	US-08-915-314-64 Sequence 64, Appl
25	76	32.9	13	4	US-08-702-054B-33 Sequence 33, Appl
26	76	32.9	14	4	US-08-915-314-57 Sequence 57, Appl
27	76	32.9	21	4	US-08-915-314-54 Sequence 54, Appl

28	75	32.5	11	4	US-08-915-314-41 Sequence 41, Appl
29	75	32.5	12	4	US-08-915-314-52 Sequence 52, Appl
30	75	32.5	13	4	US-08-702-054B-35 Sequence 35, Appl
31	75	32.5	16	4	US-08-702-054B-2 Sequence 2, Appl
32	75	32.5	18	4	US-08-702-054B-12 Sequence 12, Appl
33	74	32.0	21	4	US-08-915-314-56 Sequence 56, Appl
34	73.5	31.8	15	4	US-08-702-054B-40 Sequence 40, Appl
35	73	31.6	9	4	US-08-915-314-90 Sequence 9, Appl
36	73	31.6	11	4	US-08-915-314-44 Sequence 44, Appl
37	72	31.2	12	4	US-08-915-314-79 Sequence 79, Appl
38	72	31.2	12	4	US-08-915-314-81 Sequence 81, Appl
39	72	31.2	12	4	US-08-915-314-82 Sequence 82, Appl
40	72	31.2	12	4	US-08-915-314-84 Sequence 84, Appl
41	71	30.7	12	4	US-08-915-314-39 Sequence 39, Appl
42	71	30.7	12	4	US-08-915-314-74 Sequence 74, Appl
43	71	30.7	12	4	US-08-702-054B-5 Sequence 5, Appl
44	71	30.7	13	1	US-07-715-271-1 Sequence 1, Appl
45	71	30.7	13	1	US-08-197-205-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-915-314-46
Sequence 46, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/915,314
APPLICATION NUMBER: 660081.405
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-46

Query Match 39.2%; Score 90.5; DB 4; Length 21;
Best Local Similarity 61.5%; Pred. No. 9.1e-06;
Matches 16; Conservative 1; Mismatches 0; Indels 9; Gaps 1;
OY 1 ILRMPMPWRRKHEAPEAPIMILR 26

Db 1 ILRPMWPMRRK-----IMLK 17

RESULT 2

US-08-915-314-42
Sequence 42, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear

US-08-915-314-42

Query Match 37.2%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.8e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
Db 1 ILRPMWPMRRK 12

RESULT 3
US-08-702-054B-11
Sequence 11, Application US/08702054B
Patent No. 6191254

GENERAL INFORMATION:

APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla

STATE: CA
COUNTRY: USA
ZIP: 92037

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995

ATTORNEY/AGENT INFORMATION:

NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-054B-11

Query Match 37.2%; Score 86; DB 4; Length 16;
Best Local Similarity 41.9%; Pred. No. 2.3e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 18; Gaps 2;

OY 5 PW-WPWRKHEAPEAPIMILRPMWPMRR 34
Db 3 PWKMPW-----WPWPMRR 16

RESULT 4

US-08-915-314-47
Sequence 47, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESS: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-47

Query Match 37.2%; Score 86; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRPMWPMWRK 12
|||||
Db 1 ILRPMWPMWRK 12

RESULT 5
US-08-915-314-48

Sequence 48, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-48

Query Match 37.2%; Score 86; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 3.4e-05;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRPMWPMWRK 12
|||||
Db 1 ILRPMWPMWRK 12

RESULT 6
US-08-702-054B-38

Sequence 38, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-054B-38

Query Match 36.4%; Score 84; DB 4; Length 16;
Best Local Similarity 36.4%; Pred. No. 4.6e-05;
Matches 12; Conservative 1; Mismatches 0; Indels 20; Gaps 1;

QY 3 RMPWPMWRKHEAPEAPIMILRPMWPMWRK 35
|||||
Db 4 KMPW-----WPMWPMWRK 16

RESULT 7
US-08-915-314-40

Sequence 40, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-40

Query Match 35.9%; Score 83; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 4.3e-05;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRWPMWPMRRK 12
DB 1 ILRWPMWPMRRK 12

RESULT 8
US-08-915-314-76
Sequence 76, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 76:

SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-76

Query Match 35.5%; Score 82; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 LRWPMWPMRRK 12
DB 2 LRWPMWPMRRK 12

RESULT 9
US-08-915-314-77
Sequence 77, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-77

Query Match 35.1%; Score 81; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 8e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILRWPMWPMRRK 12
DB 1 ILRWPMWPMRRK 12

RESULT 10
US-08-915-314-87

Sequence 87, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 87:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-87

Query Match 35.1%; Score 81; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRPMWPMRR 11
Db 1 ILRPMWPMRR 11

RESULT 11
US-08-915-314-78
Sequence 78, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-78

Query Match 34.6%; Score 80; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
Db 1 ILRPMWPMRRK 12

RESULT 12
US-08-915-314-85
Sequence 85, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-85

Query Match 34.6%; Score 80; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILRPPMPWRRK 12
|||||||
Db 1 ILRPPMPWRRK 12

RESULT 13
US-08-915-314-86
Sequence 86, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-86

Query Match 34.6%; Score 80; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00011;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILRPPMPWRRK 12
|||||||
Db 1 ILRPPMPWRRK 12

RESULT 14
US-08-915-314-38
Sequence 38, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-38

Query Match 34.2%; Score 79; DB 4; Length 13;
Best Local Similarity 76.9%; Pred. No. 0.00016;
Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 23 ILRPPMPWRRK 35
|||||||
Db 1 ILRPPMPWRRK 13

RESULT 15
US-08-915-314-80
Sequence 80, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

;
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/915.314
; FILING DATE: 20-AUG-1997
;
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6180604tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 660081.405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
;
; US-08-915-314-80

```

```

Query Match 33.8%; Score 78; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.00019;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ILRWMPMPRRK 12
Db 1 ILRWMPMPRRK 12

```

Search completed: June 21, 2001, 08:38:58
Job time: 215 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:36:54 ; Search time 29.25 Seconds

(without alignments)
91.149 Million cell updates/sec

Title: SCHNIZ-444-MODESEQ2B.PEP
Perfect score: 231
Sequence: 1.ILRMPWMPWRKHEAPEPEPIMLILRMPWMPWRK 35

Scoring table: BLOSUM62
Gapop 10.0., Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	30.7	144	1 JC1222	indolicidin precursor
2	59	25.5	111	2 T29295	hypothetical prote
3	58.5	25.3	95	2 E86447	protein F5014.5 (i
4	58.5	25.3	711	2 C40046	antibiotic transpo
5	58	25.1	293	2 E83513	probable outer mem
6	57.5	24.9	485	2 S74708	hypothetical prote
7	56.5	24.5	2290	1 GNNYE	genome polyprotein
8	55	23.8	192	2 H86543	hypothetical prote
9	55	23.8	192	2 D72081	conserved hypotet
10	55	23.8	288	2 T51059	hypothetical prote
11	55	23.8	835	2 S75842	nitrogen assimilac
12	55	23.8	2292	2 S35961	capsid polyprotein
13	54.5	23.6	451	2 S30401	hypothetical prote
14	54	23.4	134	2 E72532	hypothetical prote
15	54	23.4	144	2 S35331	apidaecin 22 precu
16	54	23.4	469	2 B70607	probable cyts prot
17	54	23.4	538	2 B84759	hypothetical prote
18	53.5	23.2	352	2 S77448	hypothetical prote
19	53	22.9	107	2 T35634	hypothetical prote
20	53	22.9	449	2 C84618	hypothetical prote
21	53	22.9	1075	2 T45570	kinesin-like prote
22	53	22.9	1173	1 VG1HHC	E3 glycoprotein pr
23	53	22.9	2292	1 GNNYE	genome polyprotein
24	53	22.9	2292	1 GNNYE	capsid polyprotein
25	53	22.9	2292	1 GNNYE	hypothetical prote
26	52.5	22.7	86	2 S55401	hypothetical prote
27	52.5	22.7	173	2 F64784	ycf1 protein - Esc
28	52.5	22.7	257	2 S70177	ycf1 protein - Yec
29	52.5	22.7	361	2 A36669	galactoside 3(4)-L

30	52.5	22.7	452	2 T28094	hypothetical prote
31	52.5	22.7	478	2 A83368	hypothetical prote
32	52.5	22.7	824	2 B38423	protein-glutamine
33	52.5	22.7	1016	1 S40838	formate dehydrogen
34	52	22.5	120	2 H70817	hypothetical prote
35	52	22.5	145	2 T49102	hypothetical prote
36	52	22.5	229	2 S60454	hypothetical prote
37	52	22.5	253	2 G70715	glucose starvation
38	52	22.5	276	2 B83161	hypothetical prote
39	52	22.5	282	2 A75473	probable sigma fac
40	52	22.5	287	2 F65165	probable sigma fac
41	52	22.5	287	2 H86040	33.2 kD protein in
42	52	22.5	356	2 S74766	probable alpha hel
43	52	22.5	376	2 T26075	hypothetical prote
44	52	22.5	404	2 T05556	pectate lyase (EC
45	52	22.5	420	2 A45166	protein-lysine 6-o

ALIGNMENTS

RESULT 1
JC1222
Indolicidin precursor - bovine
N:Alternate names: antimicrobial peptide
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JC1222; A42387; S25664
R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.
Biochem. Biophys. Res. Commun. 187, 467-472, 1992
A:Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.
A:Reference number: JC1222; MUID:92392368
A:Accession: JC1222
A:Molecule type: mRNA
A:Residues: 1-144 <SAL>
A:Cross-references: EMBL:X67340; NID:9462; PIDN:CAA47755.1; PID:9463
A:Experimental source: Dore matrow
J. Biol. Chem. 267, 4292-4295, 1992
J. Biol. Chem. 267, 4292-4295, 1992
A:Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.
A:Reference number: A42387; MUID:92165771
A:Accession: A42387
A:Molecule type: Protein
A:Residues: 131-143 <SEL>
A:Experimental source: neutrophils
A:Note: sequence extracted from NCBI backbone (NCBIP:83840)
C:Superfamily: cathelin; cystatin homology
C:Keywords: amidated carboxyl end
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-129/Domain: cystatin homology <CYS>
F:131-143/Product: indolicidin #status experimental <MNT>
F:133/Modified site: amidated carboxyl end (Arg) (amide in mature form from followi

Query Match 30.7% Score 71 DB 1 Length 144
Best Local Similarity 69.2% Pred. No. 0.052
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 22 ILRMPWMPWRK 34
DB 131 ILRMPWMPWRK 143

RESULT 2
T29295
hypothetical protein C50F7.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29295
R:Johnson, D.; Steillyes, L.
submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid C50F7.

A:Reference number: Z20601
 A:Accession: T29295
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-111 <JOB>
 A:Cross-references: EMBL:U41557; PIDN:AAA83303.1; CESP:C50F7.8
 C:Genetics:
 A:Gene: CESP:C50F7.8

Query Match 25.5%; Score 59; DB 2; Length 111;
 Best Local Similarity 61.5%; Pred. No. 1.2;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 22 IMIRPMPMPRR 34
 ||:|||||
 Db 10 IMVMPMPMPGR 22

RESULT 3
 E86447
 Protein F5D14.5 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
 C:Accession: E86447
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 ansen, N.F.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewart, K.;
 Chin, C.W.; Chung, M.K.; Hultzer, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Liu, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzalli,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719
 A:Accession: E86447
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-95 <STO>
 A:Cross-references: GB:AE005172; NID:98920603; PIDN:AAE81325.1; GSPDB:GN00141
 C:Genetics:
 A:Gene: F5D14.5
 A:Map position: 1

Query Match 25.3%; Score 58.5; DB 2; Length 95;
 Best Local Similarity 21.3%; Pred. No. 1.2;
 Matches 10; Conservative 4; Mismatches 4; Indels 29; Gaps 3;

OY 4 WP-----WPPRRKHEAPEPIMIL-----RPPMPW 32
 ||:|||||
 Db 46 WPPVVVVAGVGGGRMMW-----PVLVTVDGSGSMW 81

RESULT 4
 C40046
 antibiotic transport-associated protein actIT-3 - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 22-Oct-1999
 C:Accession: C40046
 R:Fernandez-Moreno, M.A.; Caballero, J.L.; Hopwood, D.A.; Malpartida, F.
 Cell 66, 769-780, 1991
 A:Title: The act cluster contains regulatory and antibiotic export genes, direct targets
 A:Reference number: A40046; MUID:91347376
 A:Accession: C40046
 A:Molecule type: DNA
 A:Residues: 1-711 <PER>
 A:Cross-references: GB:M64683; NID:q153143; PIDN:AAA26691.1; PID:q153146

Query Match 25.3%; Score 58.5; DB 2; Length 711;
 Best Local Similarity 40.7%; Pred. No. 9.5;

Matches 11; Conservative 3; Mismatches 8; Indels 5; Gaps 1;
 OY 3 RMPMPWRKHEAPEPIMILRMPW 29
 ||:|||||
 Db 335 RMPMPWRKHEAPEPIMILRMPW 356

RESULT 5
 E83513
 Probable outer membrane protein PA1048 [imported] - Pseudomonas aeruginosa (strain PA
 C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: E83513
 R:Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; L
 Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337
 A:Accession: E83513
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-293 <STO>
 A:Cross-references: GB:AE004537; GB:AE004091; NID:99946960; PIDN:AA604437.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA1048

Query Match 25.1%; Score 58; DB 2; Length 293;
 Best Local Similarity 30.6%; Pred. No. 4.4;
 Matches 11; Conservative 6; Mismatches 9; Indels 10; Gaps 2;

OY 6 WMPW-----RRKHEAPEPIMILRMPW 32
 ||:|||||
 Db 43 WMPFKDGDGTVAKATQKAEAEHVSGSSHW-MWPF 77

RESULT 6
 S74708
 hypothetical protein slr1306 - Synechocystis sp. (strain PCC 6803)
 C:Species: Synechocystis sp.
 A:Variety: PCC 6803
 C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
 C:Accession: S74708
 R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,
 o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas
 DNA Res. 3, 109-116, 1996
 A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys
 s.
 A:Reference number: S74322; MUID:97061201
 A:Accession: S74708
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-485 <KAN>
 A:Cross-references: EMBL:D90901; GB:AB001339; NID:q1651897; PIDN:BA16859.1; PID:q101
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 24.9%; Score 57.5; DB 2; Length 485;
 Best Local Similarity 29.4%; Pred. No. 8.5;
 Matches 10; Conservative 4; Mismatches 5; Indels 15; Gaps 1;

OY 5 PW-----WPPRRKHEAPEPIM 23
 ||:|||||
 Db 41 PWDGMLMLSLGLVYWRRRRHAPPEOKML 74

RESULT 7
 GNNVE
 genome polyprotein - encephalomyocarditis virus
 N:Contains: coat protein VP1; coat protein VP2; coat protein VP3; coat protein VP4; c
 EC 3.4.-.-); RNA-directed RNA polymerase (EC 2.7.7.48)

```

C:Species: encephalomyocarditis virus, EMCV
A:Note: host Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 16-Jul-1999
C:Accession: A03906; JN0383
R:Palmenberg, A.C.; Kirby, E.M.; Janda, M.R.; Drake, N.L.; Duke, G.M.; Potratz, K.F.; CC
Nucleic Acids Res. 12, 2969-2985, 1984
A:Title: The nucleotide and deduced amino acid sequences of the encephalomyocarditis vir
A:Reference number: A03906; MUID:84169586
A:Accession: A03906
A:Molecule type: genomic RNA
A:Residues: 1-2290 <PAL>
A:Cross-references: GB:X00463; NID:961034; PIDN:CA25152.1; PID:661035
R:Petrov, N.A.; Chizhikov, V.E.; Blinov, V.M.; Karimov, V.A.; Mikryukov, N.N.; Gutorov,
Bioorg. Khim. 10, 274-279, 1984
A:Title: Nucleotide sequence of the 3'-terminus of encephalomyocarditis virus RNA.
A:Reference number: JN0383; MUID:85022788
A:Accession: JN0383
A:Molecule type: genomic RNA
A:Residues: 1337-1396, 'L', 1398-1517, 'A', 1519-1536, 'E', 1538-1556, 'S', 1558-1611, 'T', 1613-1
A:Cross-references: GB:M54935
A:Note: the authors translated the codon CAU for residue 713 as Thr and AAC for residue
C:Superfamily: foot-and-mouth disease virus genome polypeptide
C:Keywords: coat protein; core protein; genome-linked protein; hydrolase; nucleocidylttr
F:1-67/Domain: leader peptide #status predicted <DP>
F:68-136/Product: coat protein VP4 #status predicted <VP4>
F:137-391/Product: coat protein VP2 #status predicted <VP2>
F:392-622/Product: coat protein VP3 #status predicted <VP3>
F:623-910/Product: coat protein VP1 #status predicted <VP1>
F:911-1056/Product: core protein P2-A #status predicted <P2A>
F:1057-1192/Product: core protein P2-B #status predicted <P2B>
F:1193-1517/Product: core protein P2-C #status predicted <P2C>
F:1518-1605/Product: core protein P3-A #status predicted <P3A>
F:1606-1625/Product: genome-linked protein VPg #status predicted <VPG>
F:1626-1830/Product: proteinase #status predicted <PTS>
F:1831-2290/Product: RNA-directed RNA polymerase #status predicted <RDP>

Query Match 24.5%; Score 56.5; DB 1; Length 2290;
Best Local Similarity 29.2%; Pred. No. 55;
Matches 14; Conservative 5; Mismatches 14; Indels 15; Gaps 2;

OY 3 RMPWPRRKHEAPEAPIMI-----LRMPW---WPMRRK 35
D 965 RAMPNPKNTYQAVLAEPCTMDIYKRVPRFLPVOKEMVREE 1012

RESULT 8
hypothetical protein CPJ0426 [imported] - Chlamydomonada pneumoniae (strain J138)
C:Species: Chlamydomonada pneumoniae, Chlamydia pneumoniae
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Mar-2001
C:Accession: H86543
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; IS
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
A:Reference number: A86491; MUID:20330349
A:Accession: H86543
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <STO>
A:Cross-references: GB:BA000008; NID:98978798; PIDN:BA98634.1; GSPDB:GN00142
C:Genetics:
A:Gene: CPJ0426

Query Match 23.8%; Score 55; DB 2; Length 192;
Best Local Similarity 45.5%; Pred. No. 6.7;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 4 WPMWPRRKHEAPEAPIMIL 25
D 138 WPMWPRRKHEAPEAPIMIL 159


```

```

RESULT 9
D72081
Conserved hypothetical protein frameshifted CP0327 [imported] - Chlamydomonada pneum
N:Alternate names: hypothetical protein CT277 homolog
C:Species: Chlamydomonada pneumoniae, Chlamydia pneumoniae
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: D72081; G81589
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lamme, C.; Fan, J.; Olinger, L.; Grilmoor
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
A:Reference number: A72000; MUID:99206606
A:Accession: D72081
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-192 <ARN>
A:Cross-references: GB:AE001625; GB:AE001363; NID:94376695; PIDN:AD18570.1; PID:94
A:Experimental source: strain CWL029
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hic
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Saliz
Nucleic Acids Res. 28, 1397-1406, 2000
A:Title: Genome sequences of Chlamydia trachomatis MOpn and Chlamydia pneumoniae Af
A:Reference number: A81500; MUID:20150255
A:Accession: G81589
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-192 <REA>
A:Cross-references: GB:AE002195; GB:AE002161; NID:97189246; PIDN:AA38182.1; PID:97
A:Experimental source: strain AR39, HL cells
C:Genetics:
A:Gene: CP0426; CP0327

Query Match 23.8%; Score 55; DB 2; Length 192;
Best Local Similarity 45.5%; Pred. No. 6.7;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 4 WPMWPRRKHEAPEAPIMIL 25
D 138 WPMWPRRKHEAPEAPIMIL 159

RESULT 10
T51059
hypothetical protein B12F1.110 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
C:Accession: T51059
R:Schulte, U.; Allyn, V.; Heideisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyaka
submitted to the Protein Sequence Database, July 2000
A:Reference number: 225286
A:Accession: T51059
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-288 <SCH>
A:Cross-references: EMBL:AL390091; GSPDB:GN00116; NCSP:B12F1.110
A:Experimental source: BAC clone B12F1; strain OR74A
C:Genetics:
A:Gene: NCSP:B12F1.110
A:Map position: 6
A:Insertions: 154/1

Query Match 23.8%; Score 55; DB 2; Length 288;
Best Local Similarity 50.0%; Pred. No. 10;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 13 HEAPPEAPIMILRMP 28
D 58 HEPPESHPIIVQYWP 73

RESULT 11

```


S75842 nitrogen assimilation regulatory protein - *Synechocystis* sp. (strain PCC 6803)

N:Alternate names: protein sir1329

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S75842

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res: 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

B:Reference number: S74322; MUID:97061201

A:Accession: S75842

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-835 <KAN>

A:Cross-references: EMBL:D90913; GB:AB001339; NID:G165348; PIDN:BA18301.1; PID:d101903

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: RNA polymerase sigma factor interaction domain homology

F:159-366/Domain: RNA polymerase sigma factor interaction domain homology <SF1>

Query Match 23.8%; Score 55; DB 2; Length 835;

Best Local Similarity 27.1%; Pred. No. 30;

Matches 13; Conservative 4; Mismatches 13; Indels 18; Gaps 2;

2 LRPMPWRRKHEAPEAPIMT-----LRMPWMP 32

DB 411 LRSAMPDRNLNVTGTFAPFYVAILELPGPOTRDONFALNLFMAMW-W 457

RESULT 12

S35961 capsid polypeptide precursor - encephalomyocarditis virus

C:Species: encephalomyocarditis virus, EMCV

C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999

C:Accession: S35961

R:Zimmermann, A.; Nielsen-Slitz, B.; Kruppenbacher, J.P.; Eggers, H.J.

submitted to the EMBL Data Library, July 1993

A:Description: The complete nucleotide sequence and construction of a full length cDNA

A:Reference number: S35961

A:Accession: S35961

A:Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1-2292 <ZIM>

A:Cross-references: EMBL:X74312; NID:9396509; PIDN:CAA52361.1; PID:9396510

C:Superfamily: foot-and-mouth disease virus genome polypeptide

C:Keywords: polypeptide

Query Match 23.8%; Score 55; DB 2; Length 2292;

Best Local Similarity 40.9%; Pred. No. 85;

Matches 9; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

3 RMPMPWRRKHEAPEAPIMT 24

DB 967 RAPMPWRRKHEAPEAPIMT 988

RESULT 13

S30401 hypothetical protein 2 - *Streptomyces clavuligerus* plasmid pSCL

C:Species: *Streptomyces clavuligerus*

C:Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 22-Oct-1999

C:Accession: S30401

R:Wu, X.; Roy, K.L.

J. Bacteriol. 175, 37-52, 1993

A:Title: Complete nucleotide sequence of a linear plasmid from *Streptomyces clavuligerus*

A:Reference number: S30400; MUID:93106972

A:Accession: S30401

A:Molecule type: DNA

A:Residues: 1-451 <WUX>

A:Cross-references: EMBL:X54107; NID:948758; PIDN:CAA36041.1; PID:9581632

C:Genetics: A:Genome: plasmid pSCL

A:Start codon: GTG

Query Match 23.6%; Score 54.5; DB 2; Length 451;

Best Local Similarity 25.4%; Pred. No. 19;

Matches 15; Conservative 1; Mismatches 6; Indels 37; Gaps 4;

2 LRPMPWRRKHEAPEAPIMT-----LRMPWMP 32

DB 362 LRPMPWRRKHEAPEAPIMT-----LRMPWMP 411

RESULT 14

E72532 hypothetical protein APE2233 - *Aeropyrum pernix* (strain K1)

C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000

C:Accession: E72532

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*

A:Reference number: A72450; MUID:99310339

A:Accession: E72532

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-134 <KAN>

A:Cross-references: DDB:AF000063; NID:95105654; PIDN:BA18245.1; PID:d1045031; PID:9

A:Experimental source: strain K1

C:Genetics: A:Gene: APE2233

C:Superfamily: *Aeropyrum pernix* hypothetical protein APE2233

Query Match 23.4%; Score 54; DB 2; Length 134;

Best Local Similarity 28.6%; Pred. No. 6.2;

Matches 14; Conservative 2; Mismatches 13; Indels 20; Gaps 3;

5 PMPWRRKHEAPEAPIMT-----PMPW 33

DB 80 PMPWRRKHEAPEAPIMT-----PMPW 128

RESULT 15

S35331 apidactin 22 precursor - honeybee

C:Species: *Apis mellifera* (honeybee)

C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000

C:Accession: S35331

R:Casteele-Josson, K.; Capaci, T.; Casteele, P.; Tempst, P.

EMBO J. 12, 1569-1578, 1993

A:Title: Apidactin multipeptide precursor structure: a putative mechanism for amplifi

A:Reference number: S35330; MUID:93223697

A:Accession: S35331

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-144 <CAS>

A:Cross-references: EMBL:X72576; NID:9297064; PIDN:CAA51168.1; PID:9297065

C:Superfamily: procytic acidic repetitive protein

Query Match 23.4%; Score 54; DB 2; Length 144;

Best Local Similarity 64.7%; Pred. No. 6.6;

Matches 11; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

5 PMPWRRKHEAPEAPIMT 21

DB 26 PMPWRRKHEAPEAPIMT 42

Search completed: June 21, 2001, 08:39:31
Job time: 157 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:39:04 ; Search time 17.73 Seconds

(without alignments)
67.622 Million cell updates/sec

Title: SCHNIZ-444-MODESEQ2B.PEP
Sequence: 1 ILRPMWPMRKRKHEAPEPEPILRLPMWPMRKRK 35

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 93435 segs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	71	30.7	144	1	INDC_BOVIN
2	58.5	25.3	711	1	MMIA_STRCO
3	56.5	24.5	2290	1	POLG_EMCV
4	54	23.4	144	1	AP22-APIME
5	54	23.4	469	1	SYCL_MYCTU
6	53	22.9	396	1	O45B_DROME
7	53	22.9	1173	1	VGL2_CVH22
8	53	22.9	2292	1	POLG_EMCV
9	53	22.9	2292	1	POLG_EMCV
10	52.5	22.7	173	1	YBCL_ECOLI
11	52.5	22.7	361	1	FUT3_HUMAN
12	52.5	22.7	824	1	TGIC_RAT
13	52.5	22.7	1016	1	FDOS_ECOLI
14	52	22.5	253	1	P1945_MYCTU
15	52	22.5	287	1	YICC_ECOLI
16	52	22.5	420	1	LYOX_CHICK
17	52	22.5	421	1	CDS1_ARATH
18	52	22.5	465	1	FUCO_CANFA
19	52	22.5	982	1	ENV_VILV
20	52	22.5	983	1	ENV_VILV
21	52	22.5	991	1	ENV_VILV2
22	51.5	22.3	505	1	TRPE_PSSS
23	51.5	22.3	715	1	YD5_MYCTU
24	51.5	22.3	984	1	SLX1_MOUSE
25	51	22.1	559	1	INXA_CABEL
26	51	22.1	989	1	ENV_VILV1
27	50.5	21.9	68	1	Y121-BP4
28	50.5	21.9	740	1	CATA_MYCBO
29	50.5	21.9	740	1	CATA_MYCTU
30	50	21.6	283	1	AP73-APIME
31	50	21.6	324	1	RCEM_RHOGE
32	50	21.6	325	1	RCEM_CHRVI
33	50	21.6	351	1	DESA_SPIPL

34	50	21.6	418	1	FD6C_ARATH	P46312 arabidopsi
35	50	21.6	443	1	FD6C_BRANA	P48627 brassica na
36	49.5	21.4	79	1	YVAO_BACSV	P37509 bacillus su
37	49.5	21.4	126	1	YD43_MYCTU	Q11013 mycobacteri
38	49.5	21.4	196	1	YA05_SCHPO	Q09677 schizosacch
39	49.5	21.4	413	1	YBHO_ECOLI	P75771 escherichia
40	49.5	21.4	529	1	TYRO_HUMAN	P14679 homo sapien
41	49.5	21.4	888	1	UFO_MOUSE	Q00993 mus musculu
42	49.5	21.4	1790	1	SEPA_EMEI	P78621 emeritella
43	49	21.2	168	1	AP14-APIME	Q06601 apis mellif
44	49	21.2	305	1	RCEM_RHORU	P10718 rhodospirill
45	49	21.2	492	1	ADRO_BOVIN	P08165 bos taurus

ALIGNMENTS

```

RESULT 1
ID INDC_BOVIN STANDARD; PRT; 144 AA.
AC P33046;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE INDOLICIDIN PRECURSOR.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=92392368; PubMed=1520337;
RA del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;
RT "cDNA cloning of the neutrophil bactericidal peptide indolicidin.";
RL Biochem. Biophys. Res. Commun. 187:467-472(1992).
RN [2]
RP SEQUENCE OF 131-143.
RC TISSUE=Neutrophils;
RX MEDLINE=92165771; PubMed=1537821;
RA Selsted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.,
RA Cullor J.S.;
RT "Indolicidin, a novel bactericidal tridecapeptide amide from
  neutrophils.";
RL J. Biol. Chem. 267:4292-4295(1992).
CC -!- FUNCTION: POTENT MICROBICIDAL ACTIVITY, ACTIVE AGAINST
  STAPHYLOCOCCUS AUREUS AND ESCHERICHIA COLI.
CC -!- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.
CC -!- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.
CC -!- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X67340; CAA47755.1; -
CC PIR: JCI222; JC1222.
CC PIR: A42387; A42387.
CC InterPro: IPR001894; -
CC Pfam: PF00666; Cathelicidins. 1.
CC PROSITE: PS00946; CATHELICIDINS.1; 1.
CC PROSITE: PS00947; CATHELICIDINS.2; 1.
CC Antibiotic: Amidation; Signal.
CC FT SIGNAL 1 29 POTENTIAL.
CC FT PROPEP 30 130
CC FT PEPTIDE 131 143 INDOLICIDIN.
CC MOD_RES 30 30 PYRROLIDONE CARBOXYLIC ACID (BY
  SIMILARITY).

```

FT DISULFID 85 96 BY SIMILARITY.
 FT DISULFID 107 124 BY SIMILARITY.
 FT MOD.RES 143 143 AMIDATION (G-144 PROVIDE AMIDE GROUP).
 SO SEQUENCE 144 AA; 16479 MW; E3B1CBEB55C09911 CRC64;

Query Match 30.7%; Score 71; DB 1; Length 144;
 Best Local Similarity 69.2%; Pred. No. 0.023;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 22 IMLRPMWPMWR 34
 DB 131 ILMKPMWPMWR 143

RESULT 2
 ID MMLA_STRCO STANDARD; PRT; 711 AA.

AC 053902;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN ACTII-3.
 GN ACTII-3.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxId=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-91347376; PubMed-1878971;
 RA Fernandez-Moreno M.A., Caballero J.L., Hopwood D.A., Malpartida F.;
 RT "The act cluster contains regulatory and antibiotic export genes,
 RT direct targets for translational control by the hlda tRNA gene of
 RT Streptomyces";
 RL Cell 66:769-780(1991).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE MMLP FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----

DR EMBL; M64683; AAA26691.1; -;
 DR PROSITE; P550156; SSD; 2;
 DR KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 14 34 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 199 219 POTENTIAL.
 FT TRANSMEM 225 255 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 FT TRANSMEM 313 333 POTENTIAL.
 FT TRANSMEM 369 389 POTENTIAL.
 FT TRANSMEM 516 536 POTENTIAL.
 FT TRANSMEM 540 560 POTENTIAL.
 FT TRANSMEM 573 593 POTENTIAL.
 FT TRANSMEM 623 643 POTENTIAL.
 FT TRANSMEM 645 665 POTENTIAL.
 SO SEQUENCE 711 AA; 74862 MW; A546BEDABEED1B6 CRC64;

Query Match 25.3%; Score 58.5; DB 1; Length 711;
 Best Local Similarity 40.7%; Pred. No. 3.8;
 Matches 11; Conservative 3; Mismatches 8; Indels 5; Gaps 1;

OY 3 RMPWPMWRKHEAPEEPIMLRMPW 29
 DB 335 RMPWPMWRKHEAPEEPIMLRMPW 356

RESULT 3
 ID POLG_EMCV STANDARD; PRT; 2290 AA.

AC P03304;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
 DE P2A TO P2C; P3A; GENOME-LINKED PROTEIN VP6; PICOINAIN 3C
 DE (EC 3.4.22.28) (P3C); RNA-DIRECTED RNA POLYMERASE P3D
 DE (EC 2.7.7.48)]
 OS Encephalomyocarditis virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
 OC Cardiovirus.
 OX NCBI_TaxId=12104;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-84169586; PubMed-6324136;
 RA Palmenberg A.C., Kirby E.M., Janda M.R., Drake N.L., Duke G.M.,
 RA Potratz K.F., Collett M.S.;
 RT "The nucleotide and deduced amino acid sequences of the
 RT encephalomyocarditis viral polyprotein coding region";
 RL Nucleic Acids Res. 12:2969-2985(1984).
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
 CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
 CC VP3, AND VP4.
 CC -1- PTM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@sdb-sdb.ch).
 CC -----

DR EMBL; X00463; CAA25152.1; -;
 DR PIR; A03906; GNNYE.
 DR HSSP; P12296; IMEC.
 DR MEROPS; C03.009; -;
 DR MEROPS; U29.001; -;
 DR InterPro; IPR000605; -;
 DR InterPro; IPR001205; -;
 DR InterPro; IPR001676; -;
 DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
 DR Pfam; PF00910; RNA_helicase; 1.
 DR Pfam; PF00073; rhv; 3
 KW Polypeptide; Coat protein; Core protein; Transferase;
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate.
 FT PROPEP 1 67 LEADER PEPTIDE.
 FT CHAIN 68 136 COAT PROTEIN VP2 (RHQ).
 FT CHAIN 137 391 COAT PROTEIN VP2 (BETA).
 FT CHAIN 392 622 COAT PROTEIN VP3 (GAMMA).
 FT CHAIN 623 910 COAT PROTEIN VP1 (ALPHA).
 FT CHAIN 911 1056 CORE PROTEIN P2A (G).
 FT CHAIN 1057 1192 CORE PROTEIN P2B (I).
 FT CHAIN 1193 1517 CORE PROTEIN P2C (F).
 FT CHAIN 1518 1605 CORE PROTEIN P3A.
 FT CHAIN 1606 1625 GENOME-LINKED PROTEIN VP6 (H).
 FT CHAIN 1626 1830 PICOINAIN 3C (P22).
 FT CHAIN 1831 2290 RNA-DIRECTED RNA POLYMERASE P3D (E).
 FT LIPID 68 68 MYRISTATE (BY SIMILARITY).
 FT ACT SITE 1784 1784 PROTEASE (POTENTIAL).
 FT ACT SITE 1802 1802 PROTEASE (POTENTIAL).
 SO SEQUENCE 2290 AA; 255756 MW; 26BC81B7CF68CB5 CRC64;

Query Match 24.5%; Score 56.5; DB 1; Length 2290;
 Best Local Similarity 29.2%; Pred. No. 21;

Matches 14; Conservative 5; Mismatches 14; Indels 15; Gaps 2;

OY 3 RMPMPRRKHEAPEAPIMI-----LRMP-----WMPRRK 35
DB 965 RAMPNPMKNTYQAVLRAPECHVTMDIYKRVPRPLPLVQKEMVREE 1012

RESULT 4

AP22_APIME STANDARD: PRT: 144 AA.

AC P35581; P11525; P11526;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE APIDAECIN PRECURSOR, TYPE 22.

OS Apis mellifera (Honeybee).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apoecrita; Aculeata;

OC Apoidea; Apidae; Apis.

OC NCBI_TaxID=7460;

OX NCBI_TaxID=7460;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=93223697; PubMed=8467807;

RA Casteels-Josson K., Capaci T., Casteels P., Tempst P.;

RT "Apidaecin multipetide precursor structure: a putative mechanism for

amplification of the insect antibacterial response.;"

RL EMBO J. 12:1569-1578(1993).

RN [2]

RP SEQUENCE (APIDAECIN IA/IB).

RC TISSUE=Hemolymph;

RA Casteels P., Ampe C., Jacobs F., Vaeck M., Tempst P.;

RT "Apidaecins: antibacterial peptides from honeybees.;"

RL EMBO J. 8:2387-2391(1989).

-1- FUNCTION: APIDAECIN HAVE BACTERICIDAL ACTIVITY: PREDOMINANTLY

AGAINST GRAM-NEGATIVE BACTERIA. THEY SEEM TO INTERFERE WITH CELL

CC PROPAGATION.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

RESULT 5

SYCL_MYCTU STANDARD: PRT: 469 AA.

AC P66862;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE CYSTEINYL-TRNA SYNTHETASE 1 (EC 6.1.1.16) (CYSTEINE--TRNA LIGASE 1)

DE (CYRS 1).

GN CYSS1 OR CYSS OR RV3580C OR MTCY06G11.27C.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OC NCBI_TaxID=1773;

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RA MEDLINE=9634230; PubMed=9634230;

RA Cole S.T., Brosch R., Parhail J., Garner T., Churcher C., Harris D.,

Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Raverty T., Davis K., Felwell T., Gentles S., Hamlin N., Holroyd S.,

Honnsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

Rutter S., Seeger K., Skellern S., Squares S., Squares R., Sulston J.E.,

Taylor K., Whitehead S., Barrett B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

complete genome sequence.;"

RL Nature 393:537-544(1998).

-1- CATALYTIC ACTIVITY: ATP + L-CYSTEINE + TRNA(CYS) = AMP +

PYROPHOSPHATE + L-CYSTEINYL-TRNA(CYS).

-1- SUBUNIT: MONOMER (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: CYTOPLASMIC.

-1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC STRONG, TO METHIONYL-TRNA SYNTHETASE.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

RESULT 6

SYCL_MYCTU STANDARD: PRT: 396 AA.

AC P66862;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE CYSTEINYL-TRNA SYNTHETASE 1 (EC 6.1.1.16) (CYSTEINE--TRNA LIGASE 1)

DE (CYRS 1).

GN CYSS1 OR CYSS OR RV3580C OR MTCY06G11.27C.

OS Mycobacterium tuberculosis.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.

OC NCBI_TaxID=1773;

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RA MEDLINE=9634230; PubMed=9634230;

RA Cole S.T., Brosch R., Parhail J., Garner T., Churcher C., Harris D.,

Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,

Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

Raverty T., Davis K., Felwell T., Gentles S., Hamlin N., Holroyd S.,

Honnsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

Rutter S., Seeger K., Skellern S., Squares S., Squares R., Sulston J.E.,

Taylor K., Whitehead S., Barrett B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

complete genome sequence.;"

RL Nature 393:537-544(1998).

-1- CATALYTIC ACTIVITY: ATP + L-CYSTEINE + TRNA(CYS) = AMP +

PYROPHOSPHATE + L-CYSTEINYL-TRNA(CYS).

-1- SUBUNIT: MONOMER (BY SIMILARITY).

-1- SUBCELLULAR LOCATION: CYTOPLASMIC.

-1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC STRONG, TO METHIONYL-TRNA SYNTHETASE.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC


```
FT CARBOHYD 147 147 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 171 171 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 176 176 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 326 326 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 440 440 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 538 538 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 581 581 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 663 663 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 930 930 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1020 1020 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1037 1037 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1082 1082 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL)
SQ SEQUENCE 1173 AA: 128639 MW: B9CA94A1A796B3BD CRC64;
```

```
Query Match 22.98; Score 53; DB 1; Length 1173;
Best Local Similarity 62.58; Pred. No. 30;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 2 LRPMPMPV 9
Db 1112 IKPMPVW 1119
```

```
RESULT 8
POLG_EMCVB STANDARD; PRT: 2292 AA.
ID POLG_EMCVB STANDARD; PRT: 2292 AA.
AC P17593.1
DT 01-AUG-1990 (Rel. 15, Created)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C, P3A; GENOME-LINKED PROTEIN VP6; PICORNAIN 3C
DE (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)].
OS Encephalomyocarditis virus (strain emc-b nondiabetogenic).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_Taxid=12105;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89243189; PubMed=2541543;
RA Bae Y.S., Eun H.M., Yoon J.W.;
RT "Genomic differences between the diabetogenic and nondiabetogenic
RT variants of encephalomyocarditis virus."
RL Virology 170:282-287(1989).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M22457; AAA43033.1; ALT_SEQ.
CC PIR: B31473; GNMYEB.
CC HSSP: P12296; IMEC.
CC MEROPS: C03.009; -.
CC MEROPS: 029.001; -.
CC InterPro: IPR000605; -.
CC InterPro: IPR001205; -.
CC InterPro: IPR001676; -.
CC Pfam: PF00680; RNA_dep_RNA_pol; 1.
CC Pfam: PF00910; RNA_helicase; 1.
CC Pfam: PF00073; Ith; 3.
CC Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate.
FT PROPEP 1 67 LEADER PEPTIDE.
FT CHAIN 68 137 COAT PROTEIN VP4 (RHQ).
FT CHAIN 138 393 COAT PROTEIN VP2 (BETA).
FT CHAIN 394 624 COAT PROTEIN VP3 (GAMMA).
FT CHAIN 625 901 COAT PROTEIN VP1 (ALPHA).
FT CHAIN 902 1058 CORE PROTEIN P2A (G).
FT CHAIN 1059 1194 CORE PROTEIN P2B (I).
FT CHAIN 1195 1519 CORE PROTEIN P2C (F).
FT CHAIN 1520 1607 CORE PROTEIN P3A.
FT CHAIN 1608 1627 GENOME-LINKED PROTEIN VP6 (H).
FT CHAIN 1628 1832 PICORNAIN 3C (P22).
FT CHAIN 1833 2292 RNA-DIRECTED RNA POLYMERASE P3D (E).
FT LIPID 68 68 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1786 1786 PROTEASE (POTENTIAL).
FT ACT_SITE 1804 1804 PROTEASE (POTENTIAL).
SQ SEQUENCE 2292 AA: 255495 MW: 8540DDEB1437EBD4 CRC64;
```

```
Query Match 22.98; Score 53; DB 1; Length 2292;
Best Local Similarity 40.98; Pred. No. 57;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;
```

```
OY 3 RMPMPMRKHAEPEEPIMI 24
Db 967 RMPMPKNTYHAUVAEPYR 988
```

```
RESULT 9
POLG_EMCVD STANDARD; PRT: 2292 AA.
ID POLG_EMCVD STANDARD; PRT: 2292 AA.
AC P17594.1
DT 01-AUG-1990 (Rel. 15, Created)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS
DE P2A TO P2C, P3A; GENOME-LINKED PROTEIN VP6; PICORNAIN 3C
DE (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D
DE (EC 2.7.7.48)].
OS Encephalomyocarditis virus (strain emc-d diabetogenic).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Cardiovirus.
OX NCBI_Taxid=12106;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89243189; PubMed=2541543;
RA Bae Y.S., Eun H.M., Yoon J.W.;
RT "Genomic differences between the diabetogenic and nondiabetogenic
RT variants of encephalomyocarditis virus."
RL Virology 170:282-287(1989).
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAEDRAL UNITS,
CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,
CC VP3, AND VP4.
CC -1- PM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22458; AAA43034.1; -
DR PIR; A31473; GNVED.
DR HSSP; P12296; IMED.
DR MEROPS; C03.009; -
DR MEROPS; U29.001; -
DR InterPro; IPR000605; -
DR InterPro; IPR001205; -
DR InterPro; IPR001676; -
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR Pfam; PF00073; rhv; 3.
KM Polypeptide; Coat protein; Core protein; Transferase;
KM RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate.
FT PROPEP 1 67 LEADER PEPTIDE.
FT CHAIN 68 137 COAT PROTEIN VP4 (RHO).
FT CHAIN 138 393 COAT PROTEIN VP2 (BETA).
FT CHAIN 394 624 COAT PROTEIN VP3 (GAMMA).
FT CHAIN 625 901 COAT PROTEIN VP1 (ALPHA).
FT CHAIN 902 1058 CORE PROTEIN P2A (G).
FT CHAIN 1059 1194 CORE PROTEIN P2B (F).
FT CHAIN 1195 1519 CORE PROTEIN P2C (F).
FT CHAIN 1520 1607 CORE PROTEIN P3A.
FT CHAIN 1608 1627 GENOME-LINKED PROTEIN VPG (H).
FT CHAIN 1628 1832 PICORNAIN 3C (P22).
FT CHAIN 1833 2292 RNA-DIRECTED RNA POLYMERASE P2D (E).
FT LIPID 68 68 MYRISTATE (BY SIMILARITY).
FT ACT_SITE 1786 1786 PROTEASE (POTENTIAL).
FT ACT_SITE 1804 1804 PROTEASE (POTENTIAL).
SQ SEQUENCE 2292 AA; 255426 MW; F2B0627B0F44107 CRC64;

Query Match 22.94; Score 53; DB 1; Length 2292;
Best Local Similarity 40.98; Pred. No. 57;
Matches 9; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

OY 3 RMPWMPRRKHEAPEAPIMI 24
| | | | | | | | | |
DB 967 RAMPNPKNTYHAYVRAEPTRV 988

RESULT 10
YBCL_ECOLI
ID YBCL_ECOLI STANDARD; PRT; 173 AA.
AC P45570;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 19.5 KDA PROTEIN IN CYSS-FOLD INTERGENIC REGION.
GN YBCL.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
OX [1]
RN SEQUENCE FROM N.A.
RP STRAIN-K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RA Roberts D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
RA Duncan M., Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O.,
RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-144 FROM N.A.
RC STRAIN-K12 / W3110;
RA Yonekura Y., Sanpei G., Mizobuchi K.;
RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 50-173 FROM N.A.
RC STRAIN-K12;
RA Corrochano L.M.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
RN [5]
RP IDENTIFICATION.
RX MEDLINE-96032851; PubMed-7567469;
RA Borodovsky M., McIninch J., Koehn E.V., Rudd K.E., Medigue C.,
RA Danchin A.;
RT "Detection of new genes in a bacterial genome using Markov models for
RT three gene classes.";
RL Nucleic Acids Res. 23:3554-3562(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000158; AAC73629.1; -
DR EMBL; U82664; AAB40280.1; -
DR EMBL; D10588; -; NOT_ANNOTATED_CDS.
DR EMBL; X59293; -; NOT_ANNOTATED_CDS.
DR Ecogen; EG12708; ybcl.
KM Hypothetical protein.
SQ SEQUENCE 173 AA; 19530 MW; 68055C4226429356 CRC64;

Query Match 22.78; Score 52.5; DB 1; Length 173;
Best Local Similarity 24.68; Pred. No. 5.4;
Matches 14; Conservative 4; Mismatches 14; Indels 25; Gaps 4;

OY 4 WPM-----W-----PMRRKHEAPEAPIMI--LRPMP-----RRK 35
| | | | | | | | | | | | | | | |
DB 117 WPMSEDFEFAFMQYIKVAPFALSRITTPYGHQVITSELMWVFLGMLMGMLMRRR 173

RESULT 11
FUT3_HUMAN
ID FUT3_HUMAN STANDARD; PRT; 361 AA.
AC P21217; Q99448; Q99449;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GALACTOSIDE 3(4)-L-FUCOSYLTRANSFERASE (EC 2.4.1.65) (BLOOD GROUP LEWIS
DE ALPHA-4-FUCOSYLTRANSFERASE) (LEWIS FT) (FUCOSYLTRANSFERASE 3) (FUCT-
DE IIT).
GN FUT3 OR LE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-91032981; PubMed-1977660;
RA Kulkowska-Iacallo J.F., Larsen R.D., Nair R.P., Lowe J.B.;
RT "A cloned human cDNA determines expression of a mouse stage-specific
RT embryonic antigen and the Lewis blood group
RT alpha(1,3/1,4)fucosyltransferase.";
RL Genes Dev. 4:1288-1303(1990).
RL

```


RN SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RC MEDLINE-95378269; PubMed-7650030;
 RA Cameron H.S., Szczepaniak D., Weston W.;
 RT "Expression of human chromosome 19p alpha(1,3)-fucosyltransferase
 RT genes in normal tissues. Alternative splicing, polyadenylation, and
 RT isoforms.";
 RL J. Biol. Chem. 270:20112-20122(1995).
 RN [3]
 RP VARIANT LE(-) MET-105.
 RX MEDLINE-94059067; PubMed-8240322;
 RA Elmgren A., Rydberg L., Larson G.;
 RT "Genotypic heterogeneity among Lewis negative individuals.";
 RL Biochem. Biophys. Res. Commun. 196:515-520(1993).
 RN [4]
 RP VARIANTS LE(-) ARC-20; SER-170 AND ALA-336.
 RX MEDLINE-94059082; PubMed-8240337;
 RA Nishihara S., Yazawa S., Iwasaki H., Nakazato M., Kudo T., Ando T.,
 RT "Alpha (1,3/1,4)fucosyltransferase (FucT-III) gene is inactivated by
 RT a single amino acid substitution in Lewis histo-blood type negative
 RT individuals.";
 RL Biochem. Biophys. Res. Commun. 196:624-631(1993).
 RN [5]
 RP VARIANTS LE(-) ARC-20 AND SER-170.
 RX MEDLINE-94033579; PubMed-8219240;
 RA Koda Y., Kimura H., Mekada E.;
 RT "Analysis of Lewis fucosyltransferase genes from the human gastric
 RT mucosa of Lewis-positive and -negative individuals.";
 RL Blood 82:2915-2919(1993).
 RN [6]
 RP VARIANTS LE(-) ARC-20 AND LYS-356.
 RX MEDLINE-94342259; PubMed-8063716;
 RA Mollifone R., Requinque I., Kelly R.J., Fletcher A., Watt J., Oriol R.,
 RT "Molecular basis for Lewis alpha(1,3/1,4)-fucosyltransferase gene
 RT deficiency (FUT3) found in Lewis-negative Indonesian pedigrees.";
 RL J. Biol. Chem. 269:20987-20994(1994).
 RN [7]
 RP VARIANT LE(-) LYS-356.
 RX MEDLINE-95050753; PubMed-7961897;
 RA Nishihara S., Natimatsu H., Iwasaki H., Yazawa S., Akamatsu S.,
 RT Ando T., Seno T., Natimatsu I.;
 RT "Molecular genetic analysis of the human Lewis histo-blood group
 RT system.";
 RL J. Biol. Chem. 269:29271-29278(1994).
 RN [8]
 RP VARIANTS LE(-) ARC-20; ARG-68; MET-105 AND LYS-356.
 RX MEDLINE-96243526; PubMed-8801770;
 RA Elmgren A., Boerjeson C., Svensson L., Rydberg L., Larson G.;
 RT "DNA sequencing and screening for point mutations in the human Lewis
 RT 'FUT3' gene enables molecular genotyping of the human Lewis blood
 RT group system.";
 RL Vox Sang. 70:97-103(1996).
 RN [9]
 RP VARIANTS LE(-) ARG-68 AND MET-105.
 RX MEDLINE-97413801; PubMed-9268337;
 RA Elmgren A., Mollifone R., Costache M., Boerjeson C., Oriol R.,
 RT Harrington J., Larson G.;
 RT "Significance of individual point mutations, T202C and C314T, in the
 RT human Lewis 'FUT3' gene for expression of Lewis antigens by the human
 RT alpha(1,3/1,4)-fucosyltransferase, Fuc-III.";
 RL J. Biol. Chem. 272:21994-21998(1997).
 RN [10]
 RP VARIANTS LE(+) K-102; A-124 AND VARIANTS LE(-) N-162; R-223; M-270.
 RX MEDLINE-9836689; PubMed-9703429;
 RA Pang H., Liu Y., Koda Y., Soejima M., Jia J., Schlaphoff T.,
 RT "Five novel missense mutations of the Lewis gene 'FUT3' in African
 RT 'Xhosa' and Caucasian populations in South Africa.";
 RL Hum. Genet. 102:675-680(1998).
 CC -1- FUNCTION: MAY CATALYZE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES

CC INVOLVED IN THE EXPRESSION OF VIM-2, LEWIS A, LEWIS B, SIALYL,
 CC LEWIS X AND LEWIS X/SSA-1 ANTIGENS. MAY BE INVOLVED IN BLOOD
 CC GROUP LEWIS DETERMINATION; LEWIS-POSITIVE (LE(+)) INDIVIDUALS
 CC HAVE AN ACTIVE ENZYME WHILE LEWIS-NEGATIVE (LE(-)) INDIVIDUALS
 CC HAVE AN INACTIVE ENZYME.
 CC -1- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,3-BETA-D-GALACTOSYL-
 CC N-ACETYL-D-GLUCOSAMINYL-R -> GDP + 1,3-BETA-D-GALACTOSYL-
 CC (ALPHA-1,4-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.
 CC -1- PATHWAY: GLYCOSYLATION.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
 CC FORM IN TRANS CISTERNAE OF GOLGI.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN STOMACH, COLON, SMALL
 CC INTESTINE, LUNG AND KIDNEY AND TO A LESSER EXTENT IN SALIVARY
 CC GLAND, BLADDER, UTERUS AND LIVER.
 CC -1- MISCELLANEOUS: ALSO ACTS ON THE CORRESPONDING 1,4-GALACTOSYL
 CC DERIVATIVE, FORMING 1,3-L-FUCOSYL LINKS.
 CC -1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
 CC GLYCOSYLTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC EMBL: X53578; CAA37641.1; -
 CC DR EMBL: U27328; AAC50187.1; -
 CC DR EMBL: U27326; AAC50185.1; -
 CC DR EMBL: U27327; AAC50186.1; -
 CC DR EMBL: D89324; BAA13941.1; -
 CC DR EMBL: D89325; BAA13942.1; -
 CC DR PIR: A36669; A36669.
 CC MIM: 111100; -
 CC DR InterPro: IPR001503; -
 CC DR Pfam: PF00852; Fucosyl1_transf. 1.
 CC KW Transferase; Glycosyltransferase; Polymorphism; Blood group antigen.
 CC KW Signal-anchor; Golgi stack;
 CC FT DOMAIN 1 15
 CC FT TRANSMEM 16 34
 CC FT DOMAIN 35 361
 CC FT CARBOHYD 134 154
 CC FT CARBOHYD 185 185
 CC FT VARIANT 20 20
 CC FT VARIANT 68 68
 CC FT VARIANT 102 102
 CC FT VARIANT 105 105
 CC FT VARIANT 124 124
 CC FT VARIANT 162 162
 CC FT VARIANT 170 170
 CC FT VARIANT 223 223
 CC FT VARIANT 270 270
 CC FT VARIANT 336 336
 CC FT VARIANT 356 356
 CC FT
 CC SQ SEQUENCE 361 AA; 42117 MW; BFA398044F19C284 CRC64;
 CC /FTID-VAR_003430.
 CC /FTID-VAR_003429.
 CC I -> K (IN LE(-)); LESS THAN 10% REDUCTION
 CC IN ACTIVITY.
 CC /FTID-VAR_007962.
 CC G -> S (IN LE(-)); COMPLETELY INACTIVE).
 CC /FTID-VAR_003428.
 CC G -> R (IN LE(-)).
 CC /FTID-VAR_007966.
 CC V -> M (IN LE(-)).
 CC /FTID-VAR_007964.
 CC D -> A (IN LE(-)).
 CC /FTID-VAR_003429.
 CC I -> K (IN LE(-)).
 CC IN ACTIVITY.
 CC /FTID-VAR_003430.
 CC BFA398044F19C284 CRC64;

Query Match 22.7%; Score 52.5; DB 1; Length 361;
 Best Local Similarity 17.7%; Pred. No. 11;
 Matches 11; Conservative 4; Mismatches 10; Indels 37; Gaps 1;

```

OY 5 PMPWRR-----KHEAPEAPIMLW 27
   1 1 1 1 1 1
Db 9 PMPWRRCLALLFOLLVAVCFPSYLRVSRDATGSPRAPSGSSRDTPTPTLLILM 68
OY 28 PW 29
   1
Db 69 TW 70

RESULT 12
TGK RAT STANDARD; PRT; 824 AA.
AC P23606;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROTEIN-GLUTAMINE GAMMA-GLUTAMYLTRANSFERASE K (EC 2.3.2.13)
DE (TRANSGLUTAMINASE K) (TGASE K) (TCK) (EPIDERMAL TGASE).
GN TGM1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxId=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91067700; PubMed-1979171;
RA Phillips M.A., Stewart B.E., Qin Q., Chakravarty R., Floyd E.E.,
RA Jettan A.M., Rice R.H.;
RT "Primary structure of keratinocyte transglutaminase.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:9333-9337(1990).
CC -1- FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE
CC CONJUGATION OF POLYAMINES TO PROTEINS. RESPONSIBLE FOR CROSS-
CC LINKING EPIDERMAL PROTEINS DURING FORMATION OF THE STRATUM
CC CORNEUM.
CC -1- CATALYTIC ACTIVITY: PROTEIN GLUTAMINE + ALKYLAMINE -> PROTEIN
CC N(5)-ALKYLGLUTAMINE + NH(3).
CC -1- COFACTOR: CALCIUM.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE TRANSGLUTAMINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M57263; AAA63495.1; -
CC PIR; B38423; B38423.
CC HSSP; P00488; 1FIE.
CC InterPro; IPR001102; -
CC InterPro; IPR002931; -
CC Pfam; PF01841; Transglutn.core.1.
CC Pfam; PF00927; Transglutamin.C.1.
CC Pfam; PF00868; Transglutamin.N.1.
CC PROSITE; PS00547; TRANSGLUTAMINASES; 1.
CC Transferrase; Acyltransferase; Calcium-binding; Membrane.
CC ACT SITE 385
CC BY SIMILARITY.
CC SEQUENCE 824 AA; 90769 MW; A7D81C148CFD938 CRC64.

Query Match 22.7%; Score 52.5; DB 1; Length 824;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 11; Conservative 0; Mismatches 8; Indels 3; Gaps 1;
OY 3 RW---PMPWRRKHEAPEAP 21
   1 1 1 1 1 1
Db 10 RWGRSFMCPTTPSPPEPEP 31

RESULT 13

```

```

FD0G_ECOLI STANDARD; PRT; 1016 AA.
ID FD0G_ECOLI;
AC P32176; P78131;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE FORMATE DEHYDROGENASE-O MAJOR SUBUNIT (EC 1.2.1.2) (FORMATE
DE DEHYDROGENASE-O ALPHA SUBUNIT) (FDH-Z ALPHA SUBUNIT) (AEROBIC FORMATE
DE DEHYDROGENASE MAJOR SUBUNIT).
GN FD0G.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE-93347969; PubMed-8346018;
RA Plunkett G. III, Burland V.D., Daniels D.L., Blattner F.R.;
RT "Analysis of the Escherichia coli genome. III. DNA sequence of the
RT region from 87.2 to 89.2 minutes.";
RT Nucleic Acids Res. 21:3391-3398(1993).
RN [2]
RP REVISIONS TO 252-261, 344-348 AND 822.
RX STRAIN-K12 / MG1655;
RX MEDLINE-9742617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 1-190 FROM N.A., AND CHARACTERIZATION.
RX STRAIN-K12;
RX MEDLINE-9609298; PubMed-8522521;
RA Abalou H., Pommer J., Giordano G., Mandrand-Berthelot M.-A.;
RT "Expression and characterization of the Escherichia coli fdx locus
RT and a possible physiological role for aerobic formate
RT dehydrogenase.";
RT J. Bacteriol. 177:7141-7149(1995).
RN [4]
RP "FUNCTION: ALLOWS TO USE FORMATE AS MAJOR ELECTRON DONOR DURING
RP AEROBIC RESPIRATION. SUBUNIT ALPHA POSSIBLY FORMS THE ACTIVE
RP SITE.
RP -1- CATALYTIC ACTIVITY: FORMATE + NAD(+) -> CO(2) + NADH.
RP -1- COFACTOR: MOLYBDENUM (MOLYBDOPTERIN) AND SELENOCYSTEINE. THE
RP ACTIVE-SITE SELENOCYSTEINE IS ENCODED BY THE OPAL CODON, UGA.
RP MAY BIND A 4FE-4S CLUSTER.
RP -1- SUBUNIT: FORMATE DEHYDROGENASE IS A MEMBRANE-BOUND COMPLEX, FORMED
RP BY SUBUNITS ALPHA, BETA AND GAMMA.
RP -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
RP -1- SIMILARITY: BELONGS TO THE PROKARYOTIC MOLYBDOPTERIN-CONTAINING
RP OXIDOREDUCTASE FAMILY.
RP -----
RP This SWISS-PROT entry is copyright. It is produced through a collaboration
RP between the Swiss Institute of Bioinformatics and the EMBL outstation -
RP the European Bioinformatics Institute. There are no restrictions on its
RP use by non-profit institutions as long as its content is in no way
RP modified and this statement is not removed. Usage by and for commercial
RP entities requires a license agreement (see http://www.isb-sib.ch/announce/
RP or send an email to license@isb-sib.ch).
RP -----
CC EMBL; L19201; AAB03027.1; -
CC EMBL; AE000464; AAD13456.1; ALT_SEQ.
CC EMBL; X87583; CA60887.1; -
CC HSSP; P07658; 1FDI.
CC Ecogen; EGI1858; fdog.
CC InterPro; IPR001467; -
CC Pfam; PF01568; Molybdop-binding.1.
CC PROSITE; PS00551; MOLYBDOPTERIN_PROK_1; 1.
CC PROSITE; PS00490; MOLYBDOPTERIN_PROK_2; FALSE_NEG.
CC PROSITE; PS00932; MOLYBDOPTERIN_PROK_3; 1.
CC Oxidoreductase; Molybdenum; Selenocysteine; Selenum; NAD.

```

KM Iron-sulfur; 4Fe-4S.
 FT METAL 50 50 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 53 53 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 57 57 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 92 92 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT SE_CYS 196 196
 FT CONFLICT 252 261 GAKLVIDPR -> RREDCCDRSC (IN REF. 1).
 FT CONFLICT 344 348 ENGFA -> GKRLR (IN REF. 1).
 FT CONFLICT 822 822 G -> P (IN REF. 1).
 SQ SEQUENCE 1016 AA; 112502 MW; 95C06BD9633C0A7C CRC64;

Query Match 22.7% Score 52.5; DB 1; Length 1016;
 Best Local Similarity 29.5%; Pred. No. 30;
 Matches 13; Conservative 5; Mismatches 13; Indels 13; Gaps 4;

OY 2 LKMPW-WPWR-----KHEAPEEAP-----IMLRW---PMPW 32
 DB 766 LGMWAMPLNRLILYNRASADPGNPNWDPKROLKMKDKTKWGSW 809

RESULT 14
 Y945_MYCTU STANDARD; PRT; 253 AA.
 AC P71564;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE OXIDOREDUCTASE RV0945 (EC 1.-.-.-).
 GN RV0945 OR MTCY10D7.29C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV:
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsbury T., Jagsels K., Kirogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 CC -i- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 CC EMBL: 279700; CAB02005.1; -
 DR Tuberculin; RV0945; -
 DR InterPro: IPR002198; -
 DR Pfam: PF00106; adh_short; 1.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Oxidoreductase.
 FT ACT_SITE 159 159 BY SIMILARITY.
 SQ SEQUENCE 253 AA; 27138 MW; BAD937208842DA12 CRC64;

Query Match 22.5% Score 52; DB 1; Length 253;
 Best Local Similarity 34.8%; Pred. No. 9;
 Matches 8; Conservative 2; Mismatches 3; Indels 10; Gaps 1;

OY 5 PMPWPKRKHAEPEAPIMLRW 27
 DB 230 PMPW-----APLVRLMW 242

RESULT 15
 Y1CC_ECOLI STANDARD; PRT; 287 AA.
 ID Y1CC_ECOLI
 AC P23839;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE PROTEIN Y1CC.
 GN Y1CC.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=93315143; PubMed=7686882;
 RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
 RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
 RL genome: organizational symmetry around the origin of replication."
 RN Genomics 16:551-561(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=89313291; PubMed=2664418;
 RA Poulsen P., Andersen J.T., Jensen K.F.;
 RT "Molecular and mutational analysis of three genes preceding pyre on
 RL the Escherichia coli chromosome."
 RN Mol. Microbiol. 3:393-404(1989).
 RN [3]
 RP IMPORTANCE AT HIGH TEMPERATURE.
 RX MEDLINE=92021791; PubMed=1925027;
 RA Poulsen P., Jensen K.F.;
 RT "Three genes preceding pyre on the Escherichia coli chromosome are
 RT essential for survival and normal cell morphology in stationary
 RL culture and at high temperature."
 RL Res. Microbiol. 142:283-288(1991).
 CC -i- MISCELLANEOUS: ESSENTIAL FOR E. COLI UNDER SPECIAL GROWTH
 CC CONDITION.
 CC -i- SIMILARITY: STRONG, TO H. INFLUENZAE HI0467.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>
 CC or send an email to license@isb-slb.ch).
 CC -----
 CC EMBL: L10328; AAA61997.1; -
 DR EMBL: AE000441; AAC7668.1; -
 DR EMBL: X14235; CAA32451.1; -
 DR PIR: S06994; S06994.
 DR ECODBASE: C033.0; 6TH EDITION.
 DR ECODBASE: C033.1; 6TH EDITION.
 DR Ecogene: EG1192; Y1CC.
 FT CONFLICT 148 148 E -> Q (IN REF. 2).
 FT CONFLICT 184 189 RIVAKL -> SGREV (IN REF. 2).
 SQ SEQUENCE 287 AA; 33175 MW; 7368D35E06FCA3B9 CRC64;

Query Match 22.5% Score 52; DB 1; Length 287;
 Best Local Similarity 45.0%; Pred. No. 10;
 Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

Fri Jun 22 08:03:32 2001

schulz-444-modseq2b.pep.rsp

Page 10

Db 94 WVKMSDEGEINPVILRMP 113

Search completed: June 21, 2001, 08:44:54
Job time: 350 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:37:49 ; Search time 48.4 Seconds
(without alignments)
95.675 Million cell updates/sec

Title: SCHNIZ-444-MODESEQ2B.PEP
Perfect score: 231
Sequence: 1 LRMPMPWRKHEAPEPEPIMLRMPMPWRK 35

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: SP TREMBL.16:*
2: SP Archaea:*
3: SP Bacteria:*
4: SP Fungi:*
5: SP Invertebrate:*
6: SP Mammal:*
7: SP MHC:*
8: SP Organelle:*
9: SP Phage:*
10: SP Plant:*
11: SP Rodent:*
12: SP Unclassified:*
13: SP Vertebrate:*
14: SP Virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	74	32.0	1245	3 Q9Y7V5	Q9Y7V5 trichoderma
2	72	31.2	746	14 Q9JH31	Q9JH31 tt virus. o
3	67.5	29.2	723	14 Q9DUC4	Q9DUC4 tt virus. o
4	65.5	28.4	735	14 Q9DUC9	Q9DUC9 tt virus. o
5	65	28.1	49	14 Q9DTR80	Q9DTR80 tt virus. o
6	65	28.1	748	14 Q9DTR81	Q9DTR81 tt virus. o
7	59	25.5	111	5 Q18753	Q18753 caenorhabdi
8	58.5	25.3	95	10 Q9LON0	Q9LON0 arabidopsis
9	58.5	25.3	118	6 Q29270	Q29270 sus scrofa
10	58.5	25.3	147	11 Q61427	Q61427 mus musculu
11	58	25.1	293	2 Q91456	Q91456 pseudomonas
12	57.5	24.9	485	2 P72844	P72844 synechocyst
13	57.5	24.9	767	14 Q9QUD8	Q9QUD8 tt virus. h
14	57	24.7	766	14 Q91EY0	Q91EY0 tt virus. p
15	56.5	24.5	2292	14 Q66765	Q66765 encephalomy
16	56	24.2	98	10 Q9LWP9	Q9LWP9 oryza sativ
17	56	24.2	745	14 Q9MSV7	Q9MSV7 tt virus. d
18	56	24.2	745	14 Q9JG78	Q9JG78 tt virus. o
19	56	24.2	746	14 Q9JG80	Q9JG80 tt virus. o

20	56	24.2	1032	10 Q9ETJ9	Q9ETJ9 arabidopsis
21	55.5	24.0	338	5 Q17312	Q17312 ceratitis c
22	55	23.8	53	14 Q9W9H9	Q9W9H9 tt virus. o
23	55	23.8	192	2 Q9Z8B7	Q9Z8B7 chlamydia p
24	55	23.8	288	3 Q9P3H2	Q9P3H2 neurospora
25	55	23.8	504	2 P96143	P96143 thermocactin
26	55	23.8	835	2 P74210	P74210 synechocyst
27	55	23.6	2292	14 Q66850	Q66850 encephalomy
28	54.5	23.6	219	2 Q86442	Q86442 pseudomonas
29	54.5	23.6	451	2 Q05074	Q05074 streptomyces
30	54	23.4	57	14 Q9UG32	Q9UG32 tt virus. o
31	54	23.4	134	1 Q9Y9Q5	Q9Y9Q5 aeropyrum p
32	54	23.4	252	14 Q9IU06	Q9IU06 human immun
33	54	23.4	252	14 Q9IU04	Q9IU04 human immun
34	54	23.4	252	14 Q9IT26	Q9IT26 human immun
35	54	23.4	252	14 Q9ITX6	Q9ITX6 human immun
36	54	23.4	252	14 Q9IT09	Q9IT09 human immun
37	54	23.4	252	14 Q9IT07	Q9IT07 human immun
38	54	23.4	252	14 Q9IT05	Q9IT05 human immun
39	54	23.4	273	4 Q9NRV2	Q9NRV2 homo sapien
40	54	23.4	418	10 Q9LW14	Q9LW14 oryza sativ
41	54	23.4	538	10 Q64683	Q64683 arabidopsis
42	54	23.4	746	14 Q9WAX2	Q9WAX2 tt virus. 1
43	54	23.4	768	14 Q70810	Q70810 tt virus. 1
44	54	23.4	769	14 Q70798	Q70798 tt virus. 1
45	54	23.4	769	14 Q70802	Q70802 tt virus. 1

ALIGNMENTS

RESULT 1
ID Q9Y7V5 PRELIMINARY; PRT; 1245 AA.
AC Q9Y7V5;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
DE CONIDIOSPORE SURFACE PROTEIN.
GN CMP1.
OS Trichoderma harzianum.
OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
OX NCBI_TaxID=5544;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 32173;
RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
van Montagu M., Herrera Estrella A., Horvitz B.A.;
RT "Developmental regulation of a gene encoding a multidomain
RT conidiospore surface protein of Trichoderma, cmpl.1";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133651; CAB40845.1;
SQ SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0CDF8 CRC64;

Query Match 32.0%; Score 74; DB 3; Length 1245;
Best Local Similarity 33.3%; Pred. No. 0.37;
Matches 10; Conservative 1; Mismatches 9; Indels 10; Gaps 1;
QY 3 RMPMPWRKHEAPEPEPIMLRMPMPWRK 32
DB 1165 RQWMSWRPRGRG-----CWQMSW 1204
RESULT 2
ID Q9JH31 PRELIMINARY; PRT; 746 AA.
AC Q9JH31;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE ORF1.
OS TT virus.

```

DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORF1.
OS TP virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PT-TTV6;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PT-TTV6;
RX PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
  Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
  "Species-specific TTV viruses in humans and nonhuman primates and their
  phylogenetic relatedness";
RT Virology 277:368-378(2000).
DR EMBL; AB041957; BAB19308.1;
SO SEQUENCE 735 AA; 86132 MW; 9ED818DCBE6FA5D3 CRC64;

Query Match 28.4%; Score 65.5; DB 14; Length 735;
Best Local Similarity 31.0%; Pred. No. 2.5;
Matches 13; Conservative 3; Mismatches 7; Indels 19; Gaps 3;

OY 2 LRPWPWRRRKHAEPEAPLMLRPWPWRRR 35
ID 1 MAMPWRRRRRRRRRRR-----RMRRRRRRRRRRR 31
AC 09DT80 PRELIMINARY; PRT; 49 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORF1 (FRAGMENT).
OS TP virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TW9;
RX MEDLINE=20568739; PubMed=1118348;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
  Sai T., Sugai Y.;
  "TP virus MNAs detected in the bone marrow cells from an infected
  individual.";
RT Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL; AB050449; BAB1930.1;
FT NON TER 49
SO SEQUENCE 49 AA; 7225 MW; 1DA6F8F1AB69AA43 CRC64;

Query Match 28.1%; Score 65; DB 14; Length 49;
Best Local Similarity 34.3%; Pred. No. 0.2;
Matches 12; Conservative 3; Mismatches 6; Indels 14; Gaps 3

OY 2 LRPWPWRRRKHAEPEAPLMLRPWPWRRR 34
ID 1 MAMTW-MORRR-----RMPRRRRRRRR 23
AC 09DT81 PRELIMINARY; PRT; 748 AA.
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)

```

QY 22 IMILRPWPWRR 34
||: ||||| |

OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.

RP SEQUENCE FROM N.A.
RC TISSUE=SMALL INTESTINE;
RX MEDLINE=96327607; PubMed=8672129

RA Wintere A.K., Fredholm M., Davies M.;
 RT "Evaluation and characterization of a porcine small intestine cDNA
 RT library: analysis of 839 clones."
 RL Mamm. Genome 7:509-517(1996).
 DR EMBL: F15078; CAA23334.1;
 FT NON-TER 1 1
 RT NON-TER 118 118
 SQ SEQUENCE 118 AA; 13609 MM; FA72E05D7B681E7B CRC64;

Query Match 25.3%; Score 58.5; DB 6; Length 118;
 Best Local Similarity 34.6%; Pred. No. 3;
 Matches 9; Conservative 4; Mismatches 4; Indels 9; Gaps 2;

OY 8 PWRKHEAPEAPIMLRN-PWMP 32
 DB 15 PWRRRMS-----VMAEQWMPW 32

RESULT 10

O61427 PRELIMINARY; PRT; 147 AA.

AC 061427;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PROCOLLAGEN, TYPE I, ALPHA 1 (ALPHA 1 TYPE I COLLAGEN) (FRAGMENT).
 GN COL1 OR COL1A1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=LIVER;
 RX MEDLINE=94344105; Pubmed=8065328;
 RA Rhodes K., Rippe R.A., Umezawa A., Nehls M., Brenner D.A., Breindl M.;
 RT "DNA methylation represses the murine alpha 1(I) collagen promoter by
 RT an indirect mechanism."
 RL Mol. Cell. Biol. 14:5950-5960(1994).
 DR EMBL: X54876; CAA38657.1;
 DR MGI: 88467; Col1a1.
 DR InterPro: IPR001007;
 DR Pfam: PF00093; vvc; 1.
 DR PROSITE: PS01208; vWFC; 1.
 DR SMART: SM00214; vwc; 1.
 FT NON-TER 147 147
 SQ SEQUENCE 147 AA; 16652 MM; 9263BF0A91B4307D CRC64;

Query Match 25.3%; Score 58.5; DB 11; Length 147;
 Best Local Similarity 40.5%; Pred. No. 3.8;
 Matches 15; Conservative 0; Mismatches 11; Indels 11; Gaps 4;

OY 3 RWPMPRRKHEAPEAPIMLRN-PW-----WMPW 33
 DB 116 RWP--PWTW---PWTWTSMSMSPWAPRTWPW 147

RESULT 11
 ID 091456 PRELIMINARY; PRT; 293 AA.

AC 091456;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PROBABLE OUTER MEMBRANE PROTEIN.
 GN PA1048.
 OS Pseudomonas aeruginosa.
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PA01;
 RX MEDLINE=20437337; Pubmed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RT opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004537; AAG04437.1;
 DR InterPro: IPR001035;
 DR InterPro: IPR001145;
 DR Pfam: PF00691; OMPA; 1.
 DR PRINTS: PR01023; NAFLGMOTY.
 DR PRINTS: PR01021; OMPADOMAIN.
 SQ SEQUENCE 293 AA; 31509 MM; 35DAB21B5374BAF2 CRC64;

Query Match 25.1%; Score 58; DB 2; Length 293;
 Best Local Similarity 30.6%; Pred. No. 8.5;
 Matches 11; Conservative 6; Mismatches 9; Indels 10; Gaps 2;

OY 6 WMPW-----RRKHEAPEAPIMLRN-PWMPW 32
 DB 43 WMPGKDDGQTVAKQAEKAEHVGSW-WMPF 77

RESULT 12

ID P72844 PRELIMINARY; PRT; 485 AA.

AC P72844;
 DT 01-FEB-1997 (TREMblrel. 02, Created)
 DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, Last annotation update)
 DE HYPOTHETICAL 54.3 KDA PROTEIN.
 GN SLR1306.
 OS Synechocystis sp. (strain PCC 6803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 OX NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=97061201; Pubmed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirose M., Sugita M., Sasaki N., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K., Okumura S.,
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions."
 RL DNA Res. 3:109-116(1996).
 DR EMBL: D90901; BAA16859.1;
 DR Hypothetical protein.
 SQ SEQUENCE 485 AA; 54270 MM; 9C315C9CCEB03D80 CRC64;

Query Match 24.9%; Score 57.5; DB 2; Length 485;
 Best Local Similarity 29.4%; Pred. No. 16;
 Matches 10; Conservative 4; Mismatches 5; Indels 15; Gaps 1;

OY 5 PW-----WMPRRKHEAPEAPIM 23
 DB 41 PWDMGMLALSLGLVYWRWRRRHAPPEOKML 74

RESULT 13

ID 090UD8 PRELIMINARY; PRT; 767 AA.

AC 090UD8;
 DT 01-MAY-2000 (TREMblrel. 13, Created)

RESULT 15	
066765	
ID 066765	PRELIMINARY;
	PRT; 2292 AA

Search completed: June 21, 2001, 08:42:25
Job time: 276 sec

DR WPI: 1998-520800/44.
 XX
 PF New indolicidin peptide analogues - useful for, e.g. enhancing
 PF activity of antibiotic or overcoming tolerance, acquired resistance
 PT or inherent resistance of microorganisms
 XX
 PS Claim 15: Page 93; 105pp: English.
 XX
 CC The present sequence represents a specifically claimed cationic peptide
 CC from the present invention. The present invention describes compositions
 CC and methods for treating infection, especially bacterial infections. The
 CC compositions and methods use cationic peptides in combination with an
 CC antibiotic agent which are then administered to a patient to enhance the
 CC activity of the antibiotic agent, to overcome: (a) tolerance; (b)
 CC acquired resistance; and (c) inherent resistance. The combinations of
 CC antibiotics and cationic peptides can provide synergistic activity
 CC against a microorganism that is tolerant, inherently resistant, or has
 CC acquired resistance to an antibiotic agent. They can be used for killing
 CC e.g. bacteria, fungi, parasites and viruses.
 XX
 SO Sequence 13 AA;

Query Match 100.0%; Score 91; DB 19; Length 13;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKRPMWPMRRK 13
 1 ILKKRPMWPMRRK 13

RESULT 4
 AAW71690
 ID AAW71690 standard; Peptide: 13 AA.
 XX
 AC AAW71690;
 XX
 DT 11-JAN-1999 (first entry)
 XX
 DE Cationic peptide MB11 (MW 1879).
 XX
 KM MB11: cationic peptide; plasmid pK1; small cryptic plasmid;
 KM replication; RepA; vector; RAMP.
 XX
 OS Synthetic.
 OS
 PN WO9841636-A2.
 XX
 PD 24-SEP-1998.
 XX
 PF 16-MAR-1998; 98WO-CA00214.
 XX
 PR 14-MAR-1997; 97US-0040722.
 XX
 PA (BURI/) BURIAN J.
 PA (KAYW/) KAY W W.
 XX
 PI Burián J, Kay W;
 XX
 DR WPI: 1998-531571/45.
 XX
 PT Increasing plasmid copy number in a cell with the repA gene product
 PT - and an small cryptic plasmid ori sequence, useful for high level
 PT expression of e.g. cytokines, antigens or therapeutic proteins
 XX
 PS Example 13; Page 54; 82pp: English.
 XX
 CC MB11 is a small (mol.wt. 1879) cationic peptide. DNA encoding
 CC MB11 has been incorporated into vector przh-11, in which the
 CC replication leader (R21) sequence of RepA (see also AAW71686) is
 CC joined to 2 hpro peptides (see also AAW71692), to provide a
 CC vector for expression of MB11 in host cells. The invention

CC provides controlled replication plasmid vectors (RAMP vectors)
 CC comprising a replicated origin of a small cryptic plasmid and a
 CC gene encoding RepA. The vectors can reach very high levels of
 CC plasmid replication, but are not lethal to the host cell, and can
 CC be used to direct the high level expression of e.g. cytokines,
 CC antigens and therapeutic proteins.
 XX
 SO Sequence 13 AA;

Query Match 100.0%; Score 91; DB 19; Length 13;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKRPMWPMRRK 13
 1 ILKKRPMWPMRRK 13

RESULT 5
 AAY94495
 ID AAY94495 standard; Peptide: 13 AA.
 XX

AC AAY94495;
 XX
 DT 20-SEP-2000 (first entry)
 XX
 DE MB1-11 peptide derived from indolicidin.
 XX
 DE Cellulose binding domain; CBD; cationic peptide;
 KM MB1-11; indolicidin; bovine.
 KM
 XX
 OS Bos taurus.
 OS
 PN WO200031279-A2.
 XX
 PD 02-JUN-2000.
 XX
 PF 19-NOV-1999; 99WO-CA01107.
 XX
 PR 20-NOV-1998; 98US-0109218.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 PA
 PI Burián J, Bartfeld D;
 XX
 DR WPI: 2000-400086/34.
 XX
 PT Multi-domain fusion protein expression cassette used for high yield
 PT stable production of foreign peptide gene products -
 XX
 PS Disclosure; Page 24; 73pp: English.
 XX

CC A novel method allows the efficient production of cationic peptides in
 CC recombinant host cells. The method involves construction of a
 CC multi-domain fusion protein expression cassette comprising a promoter and
 CC a nucleic acid molecule expressed as an insoluble protein. The inclusion
 CC of anionic peptide sequences in the linker sequences neutralises the
 CC positive charge of the cationic peptide so that the charge of the
 CC fusion protein is controlled. This cassette allows high yield, stable
 CC production of the cationic peptide. Cationic peptides such as
 CC bovine indolicidin may be used as antimicrobial agents. The present
 CC sequence is the MB1-11 peptide. MB1-11 is a cationic peptide derived
 CC from modifications of indolicidin.
 XX

SO Sequence 13 AA;

Query Match 100.0%; Score 91; DB 21; Length 13;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKRPMWPMRRK 13

Db 1 ILKKPWPWPWK 13

RESULT 6
AA92795 standard; peptide; 13 AA.

AA92795;

29-AUG-2000 (first entry)

Indolicidin analogue, CP-11.

Magalnin; antimicrobial; transgenic plant; protease degradation; Rev4;
Indolicidin; protein production; reverse peptide.

Synthetic.

WO200026344-A1.

11-MAY-2000. 99WO-US25561.

29-OCT-1999; 99WO-US25561.

30-OCT-1998; 98US-0106373.

02-NOV-1998; 98US-0106537.

(INTE-) INTERLINK BIOTECHNOLOGIES LLC.
(KENT) UNIV KENTUCKY RES FOUND.

Everett NP, LI Q, Lawrence C, Davies MH;

WPI; 2000-365597/31.

Polypeptides for reducing proteolytic degradation of proteins
administered to, or produced by a plant comprise indolicidin or its
functional equivalents

Disclosure; Page 4; 50pp; English.

Indolicidin is a potent antimicrobial tridecapeptide, originally
purified from cytoplasmic granules of bovine neutrophils. CP-11 is an
analogue, which has better activity against *E. coli*, *Pseudomonas*
aeruginosa and *Candida albicans*, but reduced activity against
Staphylococcus aureus. A reverse peptide, Rev4 (AA92796) of indolicidin
was found to have increased stability against plant protease degradation.
Expression of antimicrobial peptides in transgenic plants suffers a major
limitation in that the foreign peptides are susceptible to rapid
degradation by proteases. The invention concerns reducing the extent of
protease degradation of a protein applied to, or produced by a plant by
administering indolicidin, Rev4 or a functional equivalent to the plant.
Transgenic plants expressing indolicidin and Rev4 are useful for
production of the antimicrobial peptides. Compositions containing
indolicidin and Rev4 are also useful for production of agronomically
important proteins in plants.

Sequence 13 AA;

Query Match 100.0%; Score 91; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.6e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPWPWPWRK 13
Db 1 ILKKPWPWPWK 13

RESULT 7
AA91773 standard; peptide; 13 AA.
XX

AC AA91773;
DT 06-JUN-2000 (first entry)

Amino acid sequence of cationic peptide MBI 11.

Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
multidrug resistance.

Synthetic.

WO9965506-A2.

23-DEC-1999.

14-JUN-1999; 99WO-CA00552.

12-JUN-1998; 98US-0096541.

(MICR-) MICROLOGIX BIOTECH INC.

Friedland HD, Krieger TJ, Taylor R, Ertle D, Fraser JR, West MHP;

WPI; 2000-223549/19.

Novel pharmaceutical composition containing optionally activated
polyoxalkylene-modified cationic peptides, useful for treating tumours

Disclosure; Page 14; 94pp; English.

This sequence represents a cationic peptide amino acid sequence, which
can be used in the pharmaceutical composition of the invention. The
invention relates to a pharmaceutical composition containing at least one
activated polyoxalkylene (APO)-modified cationic peptide. The
modification of peptides with APO increases their activity against tumour
cells, including those with a multidrug resistant phenotype. The
pharmaceutical composition can be used to treat tumours, specifically
lymphoma, leukemia, multiple myeloma, or tumours of breast, lung, ovary,
cervix, uterus, skin, prostate, liver and colon.

Sequence 13 AA;

Query Match 100.0%; Score 91; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.6e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPWPWPWRK 13
Db 1 ILKKPWPWPWK 13

RESULT 8
AA91774 standard; peptide; 13 AA.
ID AA91774;

AA91774;

06-JUN-2000 (first entry)

Amino acid sequence of cationic peptide MBI 11CN.

Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
leukemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
multidrug resistance.

Synthetic.
OS
XX
PN WO9965506-A2.

XX 23-DEC-1999.
 PD
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 DR
 XX
 PT Novel pharmaceutical composition containing optionally activated
 polyoxaalkylene-modified cationic peptides, useful for treating tumours
 PT
 XX
 PS Example 3; Page 14; 94pp: English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxaalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SO Sequence 13 AA;

Query Match 100.0%; Score 91; DB 21; Length 13;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
 |||||
 DB 1 ILKKPMPMPRRK 13

RESULT 9
 AAY91818
 ID AAY91818 standard; Peptide; 13 AA.
 XX
 AC AAY91818;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Amino acid sequence of cationic peptide MBI 11E1CN.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxaalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 KM
 XX
 OS Synthetic.
 OS
 PN WO965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 DR
 XX Novel pharmaceutical composition containing optionally activated

PT polyoxaalkylene-modified cationic peptides, useful for treating tumours
 PT
 XX
 PS Disclosure; Page 15; 94pp: English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxaalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 CC
 SO Sequence 13 AA;

Query Match 100.0%; Score 91; DB 21; Length 13;
 Best Local Similarity 100.0%; Pred. No. 5.6e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
 |||||
 DB 1 ILKKPMPMPRRK 13

RESULT 10
 AAY91819
 ID AAY91819 standard; Peptide; 13 AA.
 XX
 AC AAY91819;
 XX
 DT 06-JUN-2000 (first entry)
 XX
 DE Amino acid sequence of cationic peptide MBI 11E2CN.
 XX
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxaalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 KM
 XX
 OS Synthetic.
 OS
 PN WO965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99WO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX
 PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 DR
 XX Novel pharmaceutical composition containing optionally activated
 PT polyoxaalkylene-modified cationic peptides, useful for treating tumours
 PT
 XX
 PS Disclosure; Page 15; 94pp: English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxaalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

XX	sequence	13 AA;
SQ		

Query Match	100.0%	Score 91;	DB 21;	length 13;
Best Local Similarity	100.0%	Pred. No. 5	6e-07;	
Matches 13, Conservative	0;	Mismatches	0;	Indels 0;
				Gaps 0;

```
Qy 1 ILKKWPWPWRRK 13
    | | | | | | | |
Db 1 1LkKwPwPwrrk 13
```

RESULT 11

ID AAY91820 standard; Peptide; 13 AA.

AC AAY91820;

DT 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11E3CN.

KM Cationic peptide: tumour; pharmaceutical composition; cancer; treatment.
 KM leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KM multidrug resistance.

05 Synthetic.

PN W09965506-A2.

PD 23-DEC-1999.

PF 14-JUN-1999; 99WO-CA00552.

PR 12-JUN-1998; 98US-0096541.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP,

DR WPI; 2000-223549/19.

aa	Novel pharmaceutical composition containing optionally activated
pb	polyoxalkylene-modified cationic peptides, useful for treating tumours
pt	

PS Claim 1; Page 15; 94pp; English.

This sequence represents a cationic peptide amino acid sequence, which can be used in the pharmaceutical composition of the invention. The invention relates to a pharmaceutical composition containing at least one activated polyoxalkylene (APO)-modified cationic peptide. The modification of peptides with APO increases their activity against tumour cells, including those with a multidrug resistant phenotype. The pharmaceutical composition can be used to treat tumours, specifically lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary, cervix, uterus, skin, prostate, liver and colon.

SQ Sequence 13 AA;

Query Match	100.0%	Score 91;	DB 21;	Length 13;
Best local Similarity	100.0%	Pred. No. 5.6e-07;		
Matches 13; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	1 ILKKPWPWPWRK 13
Db	1 ILKKPWPWPWRK 13

RESULT 12

AAV24583
ID AAV24583 standard; peptide; 14 AA

AC AAY24583

Dr 18-AUG-1999 (first entry)

DE Indolizidin analogue #35.

AD Indolizidin; bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antidiarrhythmic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KM food; technical material.

OS Synthetic

PN W09807745-A2

PD 26-FEB-1998.

PF 21-AUG-1997; 97WO-US14779.

PR	13-JAN-1997;	97US-0034949.
23	01 APR-1996;	96US-002475A

XX
XX
MICROBIOLOGY BIOTECH INC
(WTCB-)[illegible]XX
1000-160000/15

XX
XX
Not a copy of the original

PT New indolicidin analogues with antimicrobial activity and related
PT nucleic acid - vectors, transformed cells and antibodies, also
PT conjugates with polyoxalkylene glycol and fatty acid to reduce
PT toxicity, useful therapeutically, as disinfectants etc.

PS Claim 13; Page 89; 129pp; English

AA
 CC (I)-(VII) containing up to 25 amino acids (aa): RxxXxxXB (I), BxxXxxXB
 CC (II), BbBxxXxxXB (II), BxxXxxBBb(A)nmbLBAGS (IV), BxxXxxBBb(A)nM
 CC (V), LBbBxxXxxXNRK (VI), LbKxxXxxXB (VII) and BbKxxXxxBB (VIII).
 CC Where 2 = p or V; x = hydrophobic residue, preferably W, B = basic aa,
 CC preferably R or K; Aa = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are *Leishmania*, *Typanosoma*, *Ascaris lumbricoides*, *Fasciola*
 CC hepatica, *Klebsiella pneumoniae*, *Bordetella pertussis*, *Staphylococcus*
 CC aureus, *Listeria*, *Clostridium*, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents; also they are
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surfacte disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/Kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.

Sequence	14 AA
<p> SQ Sequence 14 AA </p>	

Query Match	100.0%	Score 91;	DB 19;	Length 14;
Best Local Similarity	100.0%	Pred. No. 6.1e-07;		
Matches 13; Conservative	0;	Mismatches	0;	Gaps 0;

```
QY 1 ILKKWPWPWPWRK 13
    |||||
Db 1 1LKKWPWPWPWRK 13
```


RESULT 13
 AAY91811
 ID AAY91811 standard; Peptide: 14 AA.
 AC AAY91811;
 DT 06-JUN-2000 (first entry)
 DE Amino acid sequence of cationic peptide MBI 11D11H.
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 KW multidrug resistance.
 OS Synthetic.
 PN W09965506-A2.
 PD 23-DEC-1999.
 PF 14-JUN-1999; 99WO-CA00552.
 PR 12-JUN-1998; 98US-0096541.
 PS (MICR-) MICROLOGIX BIOTECH INC.
 PT Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHF;
 DR WPI; 2000-223549/19.
 XX Novel pharmaceutical composition containing optionally activated
 PT polyoxalkylene-modified cationic peptides, useful for treating tumours
 PT
 PS Disclosure; Page 15; 94pp; English.
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 SQ Sequence 14 AA:

Query Match 100.0%; Score 91; DB 21; Length 14;
 Best Local Similarity 100.0%; Pred. No. 6, 1e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
 DB 1 ILKKPMPMPRRK 13
 RESULT 14
 AAY24582
 ID AAY24582 standard; Peptide: 21 AA.
 AC AAY24582;
 DT 18-AUG-1999 (first entry)
 DE Indolicidin analogue #34.
 KW Indolicidin; bacterial infection; photo-oxidised solubiliser;
 KW antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KW food; technical material.

OS Synthetic.
 PN W09807745-A2.
 PD 26-FEB-1998.
 PF 21-AUG-1997; 97WO-0514779.
 PR 13-JAN-1997; 97US-0034949.
 PR 21-AUG-1996; 96US-0024754.
 XX (MICR-) MICROLOGIX BIOTECH INC.
 PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;
 DR WPI; 1998-169090/15.
 PT New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 PS Claim 13; Page 89; 129pp; English.
 CC AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VIII) containing up to 25 amino acids (aa): R₁X₂X₃X₄ (I), B₁X₂X₃X₄
 CC (II), B₁B₂X₃X₄X₅ (III), B₁X₂X₃X₄X₅B₆ (IV), B₁X₂X₃X₄X₅B₆(A₁)_n
 CC (V), L₁B₁X₂X₃X₄X₅X₆R₇ (VI), L₁R₁X₂X₃X₄X₅X₆R₇ (VII) and B₁X₂X₃X₄X₅B₆ (VIII).
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
 CC preferably R or K; A₁ = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 SQ Sequence 21 AA:

Query Match 100.0%; Score 91; DB 19; Length 21;
 Best Local Similarity 100.0%; Pred. No. 9, 1e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPRRK 13
 DB 1 ILKKPMPMPRRK 13
 RESULT 15
 AAY91806
 ID AAY91806 standard; Peptide: 21 AA.
 AC AAY91806;
 DT 06-JUN-2000 (first entry)
 DE Amino acid sequence of cationic peptide MBI 11D4CN.
 KW Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 KW leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 KW breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

KW multidrug resistance.
 XX
 OS Synthetic.
 XX
 PN MO9965506-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 14-JUN-1999; 99MO-CA00552.
 XX
 PR 12-JUN-1998; 98US-0096541.
 XX
 PA (MICR-) MICROLOGIX BIOTECH INC.
 PI Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 DR
 XX
 PT Novel pharmaceutical composition containing optionally activated
 PS polyoxyalkylene-modified cationic peptides, useful for treating tumours
 XX
 PS Disclosure; page 15; 94pp; English.
 XX
 CC This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.
 XX
 SQ Sequence 21 AA;

Query Match 100.0%; Score 91; DB 21; Length 21;
 Best Local Similarity 100.0%; Pred. No. 9.1e-07;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ILKKPMPMPRRK 13
 |||
 DB 1 ILKKPMPMPRRK 13

Search completed: June 21, 2001, 08:33:12
 Job time: 259 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:30:38 ; Search time 17.69 Seconds
(without alignments)

14.804 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKMPMPMRK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.AA:*
1: /cgn2_6/prodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/1aa/6CTUS.COMB.pep:*
6: /cgn2_6/prodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	91	100.0	13	4	US-08-915-314-30 Sequence 30, Appl
2	91	100.0	13	4	US-08-915-314-62 Sequence 62, Appl
3	91	100.0	13	4	US-08-915-314-63 Sequence 63, Appl
4	91	100.0	13	4	US-08-915-314-64 Sequence 64, Appl
5	91	100.0	14	4	US-08-915-314-57 Sequence 57, Appl
6	91	100.0	21	4	US-08-915-314-54 Sequence 54, Appl
7	87	95.6	12	4	US-08-915-314-52 Sequence 52, Appl
8	86	94.5	12	4	US-08-915-314-74 Sequence 74, Appl
9	86	94.5	12	4	US-08-702-054B-5 Sequence 5, Appl
10	86	94.5	13	4	US-08-915-314-58 Sequence 58, Appl
11	86	94.5	14	4	US-08-915-314-59 Sequence 59, Appl
12	86	94.5	15	4	US-08-702-054B-40 Sequence 40, Appl
13	86	94.5	20	4	US-08-915-314-55 Sequence 55, Appl
14	86	94.5	21	4	US-08-915-314-56 Sequence 56, Appl
15	85	93.4	12	4	US-08-915-314-69 Sequence 69, Appl
16	85	93.4	13	4	US-08-915-314-38 Sequence 38, Appl
17	85	93.4	13	4	US-08-915-314-45 Sequence 45, Appl
18	85	93.4	13	4	US-08-702-054B-1 Sequence 1, Appl
19	85	93.4	13	4	US-08-702-054B-17 Sequence 17, Appl
20	85	93.4	13	4	US-08-702-054B-32 Sequence 32, Appl
21	83	91.2	12	4	US-08-915-314-24 Sequence 24, Appl
22	83	91.2	13	4	US-08-915-314-49 Sequence 49, Appl
23	83	91.2	13	4	US-08-915-314-49 Sequence 50, Appl
24	83	91.2	13	4	US-08-915-314-51 Sequence 51, Appl
25	83	91.2	13	4	US-08-702-054B-30 Sequence 30, Appl
26	83	91.2	13	4	US-08-702-054B-31 Sequence 31, Appl
27	83	91.2	13	4	US-08-702-054B-34 Sequence 34, Appl

28	83	91.2	13	4	US-08-702-054B-35 Sequence 35, Appl
29	82	90.1	13	4	US-08-915-314-25 Sequence 25, Appl
30	82	90.1	13	4	US-08-915-314-66 Sequence 66, Appl
31	82	90.1	13	4	US-08-915-314-67 Sequence 67, Appl
32	82	90.1	13	4	US-08-702-054B-33 Sequence 33, Appl
33	81	89.0	11	4	US-08-915-314-75 Sequence 75, Appl
34	81	89.0	15	4	US-08-702-054B-39 Sequence 39, Appl
35	80	87.9	14	4	US-08-702-054B-18 Sequence 18, Appl
36	80	87.9	15	4	US-08-702-054B-41 Sequence 41, Appl
37	80	87.9	16	4	US-08-702-054B-2 Sequence 2, Appl
38	79.5	87.4	16	4	US-08-702-054B-38 Sequence 38, Appl
39	79	86.8	17	4	US-08-702-054B-42 Sequence 42, Appl
40	78	85.7	11	4	US-08-915-314-28 Sequence 28, Appl
41	78	85.7	12	4	US-08-915-314-40 Sequence 40, Appl
42	77	84.6	12	4	US-08-915-314-39 Sequence 39, Appl
43	77	84.6	12	4	US-08-702-054B-27 Sequence 27, Appl
44	76	83.5	12	4	US-08-915-314-77 Sequence 77, Appl
45	75	82.4	12	4	US-08-915-314-42 Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-915-314-30
Sequence 30, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H. P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-30
Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Caps 0;
1 ILKKMPMPMRK 13

DB 1 ILKKPMPMPWRRK 13

RESULT 2

US-08-915-314-62
Sequence 62, Application US/08915314

Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314

FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tenburg Ph.D., Carol

REGISTRATION NUMBER: 39,317

REFERENCE/DOCKET NUMBER: 660081.405

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ. ID NO: 63:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

FEATURE:

NAME/KEY: Modified-site

LOCATION: 13

OTHER INFORMATION: /note= "D-Form of Lysine"

US-08-915-314-63

Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKPMPMPWRRK 13
DB 1 ILKKPMPMPWRRK 13

RESULT 4
US-08-915-314-64
Sequence 64, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.

APPLICANT: West, Michael H.P.

APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

APPLICANT: Erfile, Douglas

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "D-Form of Isoleucine"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 13
OTHER INFORMATION: /note= "D-Form of Lysine"
US-08-915-314-64

Query Match 100.0%; Score 91; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 4.9e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRRK 13
DB 1 ILKKPMPMPRRK 13

RESULT 5
US-08-915-314-57
Sequence 57, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-57

Query Match 100.0%; Score 91; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 5.7e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRRK 13
DB 1 ILKKPMPMPRRK 13

RESULT 6
US-08-915-314-54
Sequence 54, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-54

Query Match 100.0%; Score 91; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 7.7e-07;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPMPRRK 13
DB 1 ILKKPMPMPRRK 13

RESULT 7
US-08-915-314-52

```

Sequence 52, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-52

Query Match 95.6%; Score 87; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.6e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 2 LKKPMPMPRRK 13
DB 1 LKKPMPMPRRK 12

RESULT 8
US-08-915-314-74
Sequence 74, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:

```

```

MEDMUT TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-74

Query Match 94.5%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2,2e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

QY 1 ILKKPMPWPMRR 12
|||||
Db 1 ILKKPMPWPMRR 12

RESULT 9
US-08-702-054B-5
Sequence 5, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
TITLE OF INVENTION: AND METHODS OF SCREENING FOR THE SAME
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,054B
FILING DATE: 23-AUG-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/002,687
FILING DATE: 23-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Halle, Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/013001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid

```

TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-702-054B-5

Query Match 94.5%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWRR 12
Db 1 ILKKPMPWRR 12

RESULT 10
US-08-915-314-58
Sequence 58, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
NUMBER OF SEQUENCES: 90
INFECTIONS USING ANALOGUES OF INDOLICIDIN
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 13 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-58

Query Match 94.5%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWRR 12
Db 1 ILKKPMPWRR 12

RESULT 11
US-08-915-314-59
Sequence 59, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
NUMBER OF SEQUENCES: 90
INFECTIONS USING ANALOGUES OF INDOLICIDIN
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 59:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-59

Query Match 94.5%; Score 86; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 2.5e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILKKPMPWRR 12
Db 1 ILKKPMPWRR 12

RESULT 12
US-08-702-054B-40
Sequence 40, Application US/08702054B
Patent No. 6191254
GENERAL INFORMATION:
APPLICANT: Falls, Timothy J.
APPLICANT: Hancock, Robert E. W.
APPLICANT: Gough, Monisha
TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/702,054B
 FILING DATE: 23-AUG-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/002,687
 FILING DATE: 23-AUG-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Halile, Lisa A.
 REGISTRATION NUMBER: 38,347
 REFERENCE/DOCKET NUMBER: 07420/013001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099
 INFORMATION FOR SEQ ID NO: 40:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-702-054B-40

Query Match 94.5%; Score 86; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 2.7e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKMPMPWRR 12
 DB 1 ILKKMPMPWRR 12

RESULT 13
 US-08-915-314-55
 Sequence 55, Application US/08915314
 Patent No. 6180604
 GENERAL INFORMATION:
 APPLICANT: Fraser, Janet R.
 APPLICANT: West, Michael H.P.
 APPLICANT: Krieger, Timothy J.
 APPLICANT: Taylor, Robert
 APPLICANT: Erfile, Douglas
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
 TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
 NUMBER OF SEQUENCES: 90
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,314
 FILING DATE: 20-AUG-1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6180604tenburg Ph.D., Carol
 REGISTRATION NUMBER: 39,317
 REFERENCE/DOCKET NUMBER: 660081.405
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 55:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-915-314-55

Query Match 94.5%; Score 86; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3.5e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKMPMPWRR 12
 DB 1 ILKKMPMPWRR 12

RESULT 14
 US-08-915-314-56
 Sequence 56, Application US/08915314
 Patent No. 6180604
 GENERAL INFORMATION:
 APPLICANT: Fraser, Janet R.
 APPLICANT: West, Michael H.P.
 APPLICANT: Krieger, Timothy J.
 APPLICANT: Taylor, Robert
 APPLICANT: Erfile, Douglas
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
 TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
 NUMBER OF SEQUENCES: 90
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SEED and BERRY LLP
 STREET: 6300 Columbia Center, 701 Fifth Avenue
 CITY: Seattle
 STATE: Washington
 COUNTRY: USA
 ZIP: 98104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/915,314
 FILING DATE: 20-AUG-1997
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6180604tenburg Ph.D., Carol
 REGISTRATION NUMBER: 39,317
 REFERENCE/DOCKET NUMBER: 660081.405
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ ID NO: 56:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 US-08-915-314-56

Query Match 94.5%; Score 86; DB 4; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.7e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILKKMPMPWRR 12
 DB 1 ILKKMPMPWRR 12

RESULT 15
 US-08-915-314-69
 Sequence 69, Application US/08915314
 Patent No. 6180604
 GENERAL INFORMATION:
 APPLICANT: Fraser, Janet R.
 APPLICANT: West, Michael H.P.
 APPLICANT: Krieger, Timothy J.

APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-69

Query Match 93.48; Score 85; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. NO. 3e-06;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 2 LKKPMPMPRRK 13
DB 1 LKKPMPMPRRK 12

Search completed: June 21, 2001, 08:33:33
Job time: 175 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:31:18 ; Search time 20.32 Seconds

(Without alignments)
48.734 Million cell updates/sec

Title: US-09-444-281-35

Perfect score: 91

Sequence: 1 ILKKMPMPMPRRK 13

Scoring table: BLOSUM62

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR.68.*

1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	80.2	144	1 JG1222	Indolicidin precursor
2	54	59.3	1173	1 VGIHHC	E2 glycoprotein pr
3	51	56.0	299	2 T12505	hypothetical prote
4	49	53.8	298	2 B72492	hypothetical prote
5	49	53.8	498	1 JT0751	ferredoxin--NADP+
6	49	53.8	527	2 S33068	myosin heavy chain
7	49	53.8	715	2 B70741	probable moey prot
8	49	53.8	1940	2 A59287	myosin heavy chain
9	48	52.7	111	2 T29295	hypothetical prote
10	47	52.2	114	2 T36208	hypothetical prote
11	47	51.6	248	2 S23449	NADH oxidase (H2O2
12	47	51.6	253	2 G70715	hypothetical prote
13	46.5	51.1	352	2 S77354	histidinol-phospha
14	46.5	51.1	621	2 S37664	peptidomeric polypro
15	46.5	51.1	630	2 S37663	peptidomeric polypro
16	46.5	51.1	1154	2 VGIHHC	E2 glycoprotein pr
17	46.5	51.1	1162	1 VGIHHC	E2 glycoprotein pr
18	46.5	51.1	1162	2 S07421	E2 glycoprotein pr
19	46.5	51.1	1162	2 S14939	E2 glycoprotein pr
20	46.5	51.1	1162	2 S14940	E2 glycoprotein pr
21	46	50.5	196	2 S55483	modulator of drug
22	46	50.5	617	2 T22175	hypothetical prote
23	46	50.5	623	2 T22177	hypothetical prote
24	46	50.5	1333	2 S65812	RNA-directed DNA p
25	45	49.5	273	2 S65812	monofunctional bio
26	45	49.5	276	2 B83161	probable short-cha
27	45	49.5	397	2 B70763	probable membrane
28	45	49.5	448	2 H72376	hypothetical prote
29	45	49.5	505	2 A39128	anthranilate synth

30	45	49.5	512	2 S19439	probable membrane
31	45	49.5	964	2 E70826	probable membrane
32	45	49.5	967	2 C70831	probable mmpL4 pro
33	45	49.5	968	2 F70746	probable mmpL2 pro
34	45	49.5	1108	2 A48508	cyclic-nucleotide
35	45	49.5	1225	1 S24284	E2 glycoprotein pr
36	45	49.5	1225	2 A36607	E2 glycoprotein pr
37	45	49.5	1235	1 VGIHMC	E2 glycoprotein pr
38	45	49.5	1324	1 VGIH59	E2 glycoprotein pr
39	45	49.5	1353	1 JQ2168	E2 glycoprotein pr
40	45	49.5	1361	2 S29998	surface protein -
41	45	49.5	1362	2 A37474	surface glycoprote
42	45	49.5	1363	1 VGIHMC	E2 glycoprotein pr
43	45	49.5	1363	1 VGIHOU	E2 glycoprotein pr
44	45	49.5	1363	1 VGIHFL	E2 glycoprotein pr
45	45	49.5	1363	1 VGIHLL9	E2 glycoprotein pr

ALIGNMENTS

RESULT 1

JC1222

Indolicidin precursor - bovine

N:Alternate names: antimicrobial peptide

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: JC1222; A42387; S25664

R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.

Biochem. Biophys. Res. Commun. 187, 467-472, 1992

A:Title: cDNA cloning of the neutrophil bactericidal peptide indolicidin.

A:Reference number: JC1222; MUID:92392368

A:Accession: JC1222

A:Molecule type: mRNA

A:Residues: 1-144 <SAL>

A:Cross-references: EMBL:X67340; NID:9462; PIDN:CMA47755.1; PID:9463

A:Experimental source: Bone marrow

F:Seled, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.O.; Smith, W.; Cullor, J.S.

J. Biol. Chem. 267, 4292-4295, 1992

A:Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.

A:Reference number: A42387; MUID:92165771

A:Accession: A42387

A:Molecule type: Protein

A:Residues: 131-143 <SEL>

A:Experimental source: neutrophils

A:Note: sequence extracted from NCBI backbone (NCBI:83840)

C:Superfamily: cathelin; cystatin homology

C:Keywords: amidated carboxyl end

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-139/Domain: cystatin homology <CYS>

F:131-143/Product: indolicidin #status experimental <MAT>

F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from follow)

Query Match 80.2%; Score 73; DB 1; Length 144;
Best Local Similarity 100.0%; Pred. No. 0.0027;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 KMPMPMPRR 12
DB 135 KMPMPMPRR 143

RESULT 2

VGIHHC

E2 glycoprotein precursor - human coronavirus (strain 229E)

N:Alternate names: peplomer glycoprotein; S glycoprotein; spike glycoprotein

C:Species: human coronavirus

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 16-Jun-2000

C:Accession: A34766; S05460

R:Raabe, T.; Schelle-Prinz, B.; Siedell, S.G.

J. Gen. Virol. 71, 1065-1073, 1990
A:Title: Nucleotide sequence of the gene encoding the spike glycoprotein of human coronavirus
A:Reference number: A34766; MUID:90264837
A:Accession: A34766
A:Molecule type: mRNA
A:Residues: 1-1173 <RA>
A:Cross-references: EMBL:X16816; NID:958926; PIDN:CAA34723.1; PID:958927
A:Experimental source: Strain 229E
R:Raabe, T.; Siddell, S.
Nucleic Acids Res. 11, 6387, 1989
A:Title: Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique
A:Reference number: A34038; MUID:8536667
A:Accession: S05460
A:Status: translation not shown
A:Status: type: mRNA
A:Molecule type: mRNA
A:Residues: 1159-1173 <RA2>
A:Cross-references: EMBL:X15654; NID:958921; PIDN:CAA33680.1; PID:g1334827
A:Superfamily: coronavirus E2 glycoprotein
C:Keywords: glycoprotein; transmembrane protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-1173/Product: E2 glycoprotein #status predicted <MAT>
F:1116-1138/Domain: transmembrane #status predicted <TM>
F:123,62,98,147,171,176,220,243,326,333,440,464,518,538,542,568,581,587,663,671,930,1015

Query Match	59.38;	Score 54;	DB 1;	Length 1173;
Best Local Similarity	85.78;	Pred. No. 7;		
Matches	6;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;

Qy	4	KWPWPWPW	10
Db	1113	KWPWPWPW	1119

RESULT 3
T12505
hypothetical protein DKFZp344C192.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 23-Jul-1999 #sequence_rev1sion 23-Jul-1999 #text_change 23-Jul-1999
C:Accession: T12505
R:Ansoore, W.; Wilkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A:Reference number: 217527
A:Accession: T12505
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-299 <ANS>
A:Cross-references: EMBL:AL096753
A:Experimental source: adult testis; clone DKFZp344C192
C:Genetics:
A:Note: DKFZp344C192.1

Query Match	56.0%	Score 51	DB 2	Length 299
Best Local Similarity	85.7%	Pred. No. 4.7		
Matches	6	Conservative	0	Mismatches 1; Indels 0; Gaps 0

Qy	6	PWMPWR	12
Db.	37	PWMPWSR	43

RESULT 4
B72492
hypothetical protein APE2577 - Aeropyrum pernix (strain K1)

C:\Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C:\Accession: B72492

C/Hawakabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Nishida, A.; DNA Res. 6, 83-101, 1999

A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain number: A72450; PMID:99310339

A:Accession: B72492
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-298 <KAM>
A:Cross-references: DDBJ:AP000064; NID:95105945; PIDD:BAAB1594.1; PTD:01045380; PTD:9
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2577

Query Match	53.8%	Score 49;	DB 2;	Length 298;
Best Local Similarity	60.0%	Pred. No. 8.8;		
Matches	6;	Conservative	2;	Mismatches 0;
			Indels	Gaps 0;

```
OY      2 LKKWPWPWR 11
          :|: ||| ||
Db      102 IKETPWPWR 111
```

RESULT 5
 JT0751
 ferredoxin--NADP+ reductase (EC 1.18.1.2), long form precursor - bovine
 N.Alternate names: adrenodoxin reductase
 C.Species: Bos primigenius taurus (cattle)
 C.Date: 14-Jul-1994 #sequence, revision 18-Oct-1996 #text change 16-Jun-2000
 C.Accession: JT0751, JT0729, J50390, S03558, P50003, A29604, S52100
 R.Rakata, Y.; Sagarra, Y.; Kono, A.; Sekimizu, K.; Horikuchi, T.
 Biol. Pharm. Bull. 16, 1200-1206, 1993
 A.Title: Gene structure of bovine adrenodoxin reductase.
 A.Reference number: JT0751; MUID:94177140

A;Accession: JT0751
A;Molecule type: DNA

A1:Accession: U01290, NCBI
 A1:Cross-references: GI:DB3475; NID:91199916; PIDD:BA11921.1; PID:4521308
 A1:Experimental source: adrenal cortex
 A1:Note: the authors translated the codon GTC for residue 205 as Gly
 R1:Sagira, Y., Takata, Y., Miyata, T., Hara, T., Horuchi, T.

J. Biochem. 102, 1333-1336, 1987
A:Title: Cloning and sequence analysis of adrenodoxin reductase cDNA from bovine adrenal
A:Reference number: JT0079; MUID:88198050
A:Accession: JT0079

A: Molecule type: mRNA
A: Residues: 1-204, 211-498 <SAG>
A: Cross-references: GB:000211; NID:9217433; PIDN:BA00150.1; PID:9217434
A: Note: the deduced sequence is partially confirmed by amino acid sequencing of 15 isoforms
R: Sagara, Y.

submitted to DDBJ, September 1989
A:Reference number: JS0390
A:Contents: revision, insertion of residues 205-210
A:Accession: JS0390
A:Molecule type: mRNA

A:Residues: 56-498 <SA2>
R:Hemukoğlu, I.; Gutfinger, T.
Eur. J. Biochem. 180, 479-484, 1989
A:Title: CDNA sequence of adrenodoxin reductase. Identification of NADP-binding sites
A:Reference number: S03558; MUID:89170752

A:Accession: S03558
A:Molecule type: mRNA
A:Residues: 155-204, 211-498 <HAN>
A:Cross-references: EMBL:X13736; NID:g65; PIDD:CAA32002.1; PID:g833776
A:Note: 405-Ser was also found

R; Hamamoto, I.; Kurokouchi, K.; Tanaka, S.; Ichikawa, Y.
 Blochim. Biophys. Acta 953, 207-213, 1998
 A; Title: Adrenoferrdoxin-binding peptide of NADPH-adrenoferrdoxin reductase
 A; Reference number: PS0003; MUID:8814054
 A; Accession: PS0003

A: Molecule type: protein
A: Residues: 33-41 / S: 43-62: 260-283, / TM: 496-498 <HAW>
A: Note: a tyrosine bromide peptide binds to adrenoreceptorin
R: Konaka, Y., Murakami, H., Yabasaki, Y., Kuramitsu, S., Kagamiyama, H., Yamano, T.,
Blochem. Biophys. Res. Commun. 145, 1239-1247, 1987
A: Title: Molecular cloning and sequence analysis of full-length cDNA for mRNA of adre
A: Reference number: A29604; MUID: 87270696

A:Accession: A29604
 A:Molecule type: mRNA
 A:Residues: 1-76; 'R', 78-80, 'VWLALTPPSRML', 95-123, 'RVYRLT', 129-204, 211-273, 'R', 275-322.
 A:Cross-references: GB:M17029; NID:g162628; PIDN:AAA30362.1; PID:g162629
 A:Experimental sources: adrenal cortex
 R:Marburton, R.J.; Seybert, D.W.
 Biochim. Biophys. Acta 1246, 39-46, 1995
 A:Title: Structural and functional characterization of bovine adrenodoxin reductase by 1
 A:Reference number: S52100; MUID:95110846
 A:Accession: S52100
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 'X', 34-41, 'X', 43-48, 'X', 50-51, 304-306, 'X', 308-309, 'X', 311-326 <NAR>
 C:Comment: Ferredoxin--NADP+ reductase is localized in the matrix of adrenal cortex mitochondria.
 ferredoxin--NADP+ reductase, adrenodoxin and two forms of cytochrome P-450.
 C:Genetics:
 A:Insertions: 27/1; 59/3; 91/3; 132/3; 170/3; 204/3; 246/3; 275/1; 341/3; 399/1; 456/1
 C:Function:
 A:Description: catalyzes the reversible reduction of NADP+ by reduced ferredoxin or reduced ferredoxin.
 C:Superfamily: human ferredoxin--NADP+ reductase
 C:Keywords: alternative splicing; flavoprotein; mitochondrion; monomer; NADP; oxidoreductase
 F:1-32/Domain: transit peptide (mitochondrion) #status predicted <SIG>
 F:33-498/Product: ferredoxin--NADP+ reductase, long form #status predicted <MAT>
 F:33-204, 211-498/Product: ferredoxin--NADP+ reductase, short form #status experimental <F:40-70/Region: beta-alpha-beta FAD nucleotide-binding fold
 F:180-190/Region: NADP binding #status predicted
 F:281/Binding site: substrate (Lys) #status experimental

Query Match 53.8%; Score 49; DB 1; Length 498;
 Best Local Similarity 83.3%; Pred. No. 14;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 5 WPMWPM 10
 Db 6 WPMWPM 11

RESULT 6
 myosin heavy chain - fluke (Schistosoma mansoni) (fragment)
 N:Alternate names: surface antigen, 200K
 C:Species: Schistosoma mansoni
 C:Date: 22-Nov-1993 #sequence_revision 06-Sep-1996 #text_change 13-Feb-1998
 C:Accession: S33068
 R:Solissson, L.M.A.; Masterson, C.P.; Tom, T.D.; McNally, M.T.; Lowell, G.H.; Strand, M.
 J. Immunol. 149, 3612-3620, 1992
 A:Title: Induction of protective immunity in mice using a 62-kDa recombinant fragment of
 A:Reference number: A46514; MUID:93056536
 A:Accession: S33068
 A:Molecule type: mRNA
 A:Residues: 1-527 <SOI>
 A:Cross-references: EMBL:X65591
 A:Note: the authors translated the codon CAA for residue 346 as Lys
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: ATP; surface antigen

Query Match 53.8%; Score 49; DB 2; Length 527;
 Best Local Similarity 62.5%; Pred. No. 15;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ILKKPWW 8
 Db 106 VLKRWPMW 113

RESULT 7
 B70741
 probable moey protein - Mycobacterium tuberculosis (strain H37RV)
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: B70741

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gord
 ; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroy
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellern, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete ge
 A:Reference number: A70500; MUID:98255987
 A:Accession: B70741
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-715 <COL>
 A:Cross-references: GB:275555; GB:AL123456; NID:g3261608; PIDN:CAA9988.1; PID:e250
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: moey

Query Match 53.8%; Score 49; DB 2; Length 715;
 Best Local Similarity 60.0%; Pred. No. 20;
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 KKPMPWPR 12
 Db 64 KKPMPWPR 73

RESULT 8
 myosin heavy chain - fluke (Schistosoma mansoni) (strain Brazilian LE)
 C:Species: Schistosoma mansoni
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 08-Sep-2000
 C:Accession: A59287
 R:Weston, D.S.; Schmitz, J.; Kemp, M.; Kunz, W.
 Mol. Biochem. Parasitol. 58, 161-164, 1993
 A:Title: Cloning and sequence characterization of a complete myosin heavy chain cDN
 A:Reference number: A59287; MUID:93211444
 A:Accession: A59287
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-1940 <WES>
 A:Cross-references: GB:L01634; PIDN:AA29905.1
 A:Experimental source: strain Brazilian LE
 C:Genetics:
 A:Gene: MYH
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 F:82-752/Domain: myosin motor domain homology <CMO>

Query Match 53.8%; Score 49; DB 2; Length 1940;
 Best Local Similarity 62.5%; Pred. No. 53;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 ILKKPWW 8
 Db 809 VLKRWPMW 816

RESULT 9
 T29295
 hypothetical protein C50F7.8 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T29295
 R:Johnson, D.; Stellyes, L.
 submitted to the EMBL Data Library, November 1995
 A:Description: The sequence of C. elegans cosmid C50F7.
 A:Reference number: 220601
 A:Accession: T29295
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-111 <COH>
 A:Cross-references: EMBL:U41557; PIDN:AAA83303.1; CESP:C50F7.8
 C:Genetics:

A:Gene: CESP:C50F7.8

Query Match 52.7%; Score 48; DB 2; Length 111;

Best Local Similarity 75.0%; Pred. NO. 4.6;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 5 WPMWPMR 12

DB 15 WPMWPMGR 22

RESULT 10

T36208

hypothetical protein SCE36_09 - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T36208

R:Oliver, R.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1999

A:Reference number: Z21601

A:Accession: T36208

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-114 <COLI>

A:Cross-references: EMBL:AL049763; PIDN:CAB42078.1; GSPDB:GN00070; SCOEDB:SCE36.09

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCE36.09

Query Match 52.2%; Score 47.5; DB 2; Length 114;

Best Local Similarity 63.6%; Pred. NO. 5.5;

Matches 7; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

OY 3 KRW-PWMPWR 12

DB 102 RMRPMPWR 112

RESULT 11

S23449

NADH oxidase (H2O2-forming) (EC 1.6.1.1) - Thermus aquaticus

C:Species: Thermus aquaticus

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 23-Mar-1993

C:Accession: S23449; S24556

R:Park, H.J.; Kreutzer, R.; Reiser, C.O.A.; Sprinzl, M.

Eur. J. Biochem. 205, 875-879, 1992

A:Title: Molecular cloning and nucleotide sequence of the gene encoding a H(2)O(2)-form

A:Reference number: S23449; MUID:92249331

A:Accession: S23449

A:Molecule type: DNA

A:Residues: 1-248 <PAR>

A:Cross-references: EMBL:X60110

A:Accession: S24556

A:Molecule type: protein

A:Residues: 1-32 <PAR>

C:Genetics:

A:Gene: nox

C:Keywords: NAD: oxidoreductase

F:1-248/Product: NADH oxidase (H2O2-forming) #status experimental <MAT>

Query Match 51.6%; Score 47; DB 2; Length 248;

Best Local Similarity 100.0%; Pred. NO. 14;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PMWPM 10

DB 179 PMWPM 183

RESULT 12

G70715 hypothetical protein RV0945 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: G70715

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garrier, T.; Churcher, C.; Harris, D.; Gordon

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: G70715

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-253 <COL>

A:Cross-references: GB:Z79700; GB:AL123456; NID:g3261628; PIDN:CAB02005.1; PID:g15242

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: RV0945

C:Superfamily: rblitol dehydrogenase; short-chain alcohol dehydrogenase homology

F:8-190/Domain: short-chain alcohol dehydrogenase homology <SADR>

Query Match 51.6%; Score 47; DB 2; Length 253;

Best Local Similarity 100.0%; Pred. NO. 14;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 PMWPM 10

DB 230 PMWPM 234

RESULT 13

S77354

histidinol-phosphate aminotransferase hsc-1 - Synecocystis sp. (strain PCC 6803)

N:Alternate names: protein sll1713

C:Species: Synecocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S77354

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys

s

A:Reference number: S74322; MUID:97061201

A:Accession: S77354

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-352 <KAN>

A:Cross-references: EMBL:D90906; GB:AB001339; NID:g1652492; PIDN:BA17457.1; PID:g165

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Genetics:

A:Gene: hsc-1

C:Superfamily: probable histidinol-phosphate transaminase

Query Match 51.1%; Score 46.5; DB 2; Length 352;

Best Local Similarity 50.0%; Pred. NO. 22;

Matches 8; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

OY 2 LKKPMW----WPMR 12

DB 106 LKTRPMWQDQPMWR 121

RESULT 14

S37664

peplomeric polyprotein precursor - avian infectious bronchitis virus (strain D1466) (

N:Contains: E2 glycoprotein subunit S2

C:Species: avian infectious bronchitis virus, IBV

A:Variety: strain D1466

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
 C:Accession: S37664
 R:Kusters, J.G.; Jager, E.J.; Niesters, H.G.M.; van der Zeijst, B.A.M.
 Vaccine 8, 605-608, 1990
 A:Title: Sequence evidence for RNA recombination in field isolates of avian coronavirus
 A:Reference number: S37663; MUID:91205880
 A:Accession: S37664
 A:Molecule type: genomic RNA
 A:Residues: 1-621 <KUS>
 A:Cross-references: EMBL:X58001; NID:958986; PIDN:CAA1065.1; PID:958987
 C:Superfamily: coronavirus E2 glycoprotein
 C:Keywords: glycoprotein; peplomer protein; spike protein
 F:1-5/Product: E2 glycoprotein subunit S1 (fragment) #status predicted <GS1>
 F:6-621/Product: E2 glycoprotein subunit S2 #status predicted <GS2>
 F:10,47,59,137,144,415,447,482,506,519,542/Binding site: carbohydrate (Asn) (covalent) #

Query Match 51.1%; Score 46.5; DB 2; Length 621;
 Best Local Similarity 61.5%; Pred. No. 38;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

OY 1 ILK---KMPMPW 10
 ||| ||||: |
 Db 553 ILKTYIKMPMYW 565

RESULT 15

S37663
 peplomeric polyprotein precursor - avian infectious bronchitis virus (strain D207) (frag
 N:Contains: E2 glycoprotein subunit S2
 C:Species: avian infectious bronchitis virus, IBV
 A:Variety: strain D207
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Sep-1999
 C:Accession: S37663
 R:Kusters, J.G.; Jager, E.J.; Niesters, H.G.M.; van der Zeijst, B.A.M.
 Vaccine 8, 605-608, 1990
 A:Title: Sequence evidence for RNA recombination in field isolates of avian coronavirus
 A:Reference number: S37663; MUID:91205880
 A:Accession: S37663
 A:Molecule type: genomic RNA
 A:Residues: 1-630 <KUS>
 A:Cross-references: EMBL:X58003; NID:958988; PIDN:CAA1067.1; PID:958989
 C:Superfamily: coronavirus E2 glycoprotein
 C:Keywords: glycoprotein; peplomer protein; spike protein
 F:1-5/Product: E2 glycoprotein subunit S1 (fragment) #status predicted <GS1>
 F:6-621/Product: E2 glycoprotein subunit S2 #status predicted <GS2>
 F:10,47,59,137,144,415,447,482,506,519,542/Binding site: carbohydrate (Asn) (covalent) #

Query Match 51.1%; Score 46.5; DB 2; Length 630;
 Best Local Similarity 61.5%; Pred. No. 39;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

OY 1 ILK---KMPMPW 10
 ||| ||||: |
 Db 553 ILKTYIKMPMYW 565

Search completed: June 21, 2001, 08:33:57
 Job time: 159 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:33:13 ; Search time 11.81 Seconds

(without alignments)
37.707 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKMPMPRRK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	73	80.2	144	1	INDC_BOVIN
2	54	59.3	1173	1	VGL2_CVH22
3	49	53.8	492	1	ADRO_BOVIN
4	49	53.8	715	1	ID55_MYCTU
5	47	51.6	253	1	Y945_MYCTU
6	46.5	51.1	1154	1	VGL2_IBVD2
7	46.5	51.1	1162	1	VGL2_IBVB
8	46.5	51.1	1162	1	VGL2_IBVK
9	46.5	51.1	1162	1	VGL2_IBVM
10	46.5	51.1	1163	1	VGL2_IBV6
11	46	50.5	196	1	YA05_SCHPO
12	45	49.5	397	1	MM6_MYCTU
13	45	49.5	505	1	TRPE_PSESS
14	45	49.5	512	1	FEN2_YEAST
15	45	49.5	964	1	MM45_MYCTU
16	45	49.5	967	1	MM45_MYCTU
17	45	49.5	968	1	MM42_MYCTU
18	45	49.5	1108	1	CN3B_RAT
19	45	49.5	1225	1	VGL2_CVPR8
20	45	49.5	1225	1	VGL2_CVPR8
21	45	49.5	1235	1	VGL2_CVPR8
22	45	49.5	1324	1	VGL2_CVPR8
23	45	49.5	1353	1	VGL2_CVPR8
24	45	49.5	1363	1	VGL2_CVPR8
25	45	49.5	1363	1	VGL2_CVPR8
26	45	49.5	1363	1	VGL2_CVPR8
27	45	49.5	1363	1	VGL2_CVPR8
28	45	49.5	1363	1	VGL2_CVPR8
29	45	49.5	1363	1	VGL2_CVPR8
30	45	49.5	1376	1	VGL2_CVPR8
31	45	49.5	1447	1	VGL2_CVPR8
32	45	49.5	1447	1	VGL2_CVPR8
33	45	49.5	1447	1	VGL2_CVPR8

34	45	49.5	1447	1	VGL2_CVPR8	001977 porcine tra
35	45	49.5	1449	1	VGL2_CVPR8	P18450 porcine tra
36	45	49.5	1449	1	VGL2_CVPR8	P33470 porcine tra
37	45	49.5	1451	1	VGL2_CVPR8	P36300 canine ente
38	45	49.5	1452	1	VGL2_CVPR8	P10033 feline inte
39	45	49.5	2116	1	MY52_DICD1	P08799 dictyostell
40	44	48.4	151	1	Y0H3_PLA5	P14589 plasmidum
41	44	48.4	151	1	Y0H3_PLA5	P75709 escherichia
42	44	48.4	361	1	FUT3_HUMAN	P21217 homo sapien
43	44	48.4	372	1	FUT3_PANTR	P01058 pan troglod
44	44	48.4	451	1	MEME_ECOLI	P37353 escherichia
45	44	48.4	535	1	YDM6_SCHPO	O13912 schizosacch

ALIGNMENTS

RESULT	ID	INDC_BOVIN	STANDARD	PRT	144 AA
AC	P33046				
DR	01-OCT-1993	(Rel. 27, Created)			
DT	01-OCT-1993	(Rel. 27, Last sequence update)			
DT	01-NOV-1997	(Rel. 35, Last annotation update)			
DE	INDOLICIDIN PRECURSOR.				
OS	Bos taurus (Bovine).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;				
OX	NCHI_TaxID=9913;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Bone marrow;				
RX	MEDLINE=92392368; PubMed=1520337;				
RA	del Sal G., Storici P., Schneider C., Romeo D., Zanetti M.;				
RT	"cdna cloning of the neutrophil bactericidal peptide indolicidin.";				
RL	Biochem. Biophys. Res. Commun. 187:467-472(1992).				
RN	[2]				
RP	SEQUENCE OF 131-143.				
RC	TISSUE-Neutrophils;				
RX	MEDLINE=92165771; PubMed=1537821;				
RA	Seisted M.E., Novotny M.J., Morris W.L., Tang Y.-Q., Smith W.;				
RT	"Indolicidin, a novel bactericidal tridecapeptide amide from				
RL	neutrophils.";				
RN	J. Biol. Chem. 267:4292-4295(1992).				
CC	-1- FUNCTION: POTENT MICROBICIDAL ACTIVITY. ACTIVE AGAINST				
CC	STRAPHLOCCUS AUREUS AND ESCHERICHIA COLI.				
CC	-1- TISSUE SPECIFICITY: LARGE GRANULES OF NEUTROPHILS.				
CC	-1- PTM: ELASTASE MIGHT BE RESPONSIBLE FOR ITS MATURATION.				
CC	-1- SIMILARITY: BELONGS TO THE CATHELICIDIN FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation-				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@sib-sib.ch).				
CC	-----				
CC	EMBL: X67340; CAA47755.1; -				
CC	PIR: JCI222; JCI222.				
CC	PIR: A42387; A42387.				
CC	InterPro: IPR001894; -				
CC	Pfam: PF00666; Cathelicidins; 1.				
CC	PROSITE: PS00946; Cathelicidins; 1.				
CC	PROSITE: PS00947; Cathelicidins; 2; 1.				
CC	Antibiotic: Amidation; Signal.				
CC	SIGNAL				
CC	PROPEP				
CC	PEPTIDE				
CC	MOD_RES				
CC	INDOLICIDIN.				
CC	PYROLIDONE CARBOXYLIC ACID (BY				
CC	SIMILARITY).				

[illegible]

Query Match	Best Local Similarity	Score 54;	DB 1;	Length 1173;
Matches	6; Conservative	85.7%;	Pred. No. 2.4;	
		0; Mismatches	1; Indels	0; Gaps
Oy	4 KWPMPW 10			
Db	1113 KWPMPW 1119			
RESULT 3				
ADRO_BOVIN	STANDARD:	PRT:	492 AA.	
AC	P08165;			
DT	01-AUG-1988 (Rel. 08, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	01-OCT-2000 (Rel. 40, Last annotation update)			
DE	NADPH:ADRENODOXIN OXIDOREDUCTASE, MITOCHONDRIAL PRECURSOR			
DE	(EC 1.18.1.2) (ADRENODOXIN REDUCTASE) (AR) (FERREDOXIN-NADP(+)			
DE	REDUCTASE).			
GN	FDXR OR ADXR.			
OS	Bos taurus (bovine);			
OC	Eumariophyta, Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.			
RX	MEDLINE=94177140; PubMed=8130767.			
RA	Takata Y., Sagara Y., Kono A., Sekimizu K., Horiuchi T.;			
RT	"Gene structure of bovine adrenodoxin reductase."			
RL	Biol. Pharm. Bull. 16:1200-1206(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.			
RX	MEDLINE=86198050; PubMed=3448086;			
RA	Sagara Y., Takata Y., Miyata T., Hara T., Horiuchi T.;			
RT	"Cloning and sequence analysis of adrenodoxin reductase cDNA from			
RL	bovine adrenal cortex."			
RL	J. Biochem. 102:1333-1336(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87270696; PubMed=3038094;			
RA	Nonaka Y., Murakami H., Yabusaki Y., Kuramitsu S., Kagamiyama H.;			
RT	"Molecular cloning and sequence analysis of full-length cDNA for mRNA			
RL	of adrenodoxin oxidoreductase from bovine adrenal cortex."			
RL	Biochem. Biophys. Res. Commun. 145:1239-1247(1987).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Adrenal cortex;			
RX	MEDLINE=891170752; PubMed=2924777;			
RT	Hanukoglu I., Gutfinger T.;			
RT	"cDNA sequence of adrenodoxin reductase. Identification of NADP-			

RN binding sites in oxidoreductases.";
 RL Eur. J. Biochem. 180:479-484(1989).
 [5]
 RN SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.
 RP TISSUE-Adrenal cortex;
 RX MEDLINE-88082777; PubMed-3691502;
 RA Hanukoglu I., Gutfinger T., Hanlu M., Shively J.E.;
 RT "Isolation of a cDNA for adrenodoxin reductase (ferredoxin-NADP+
 RL reductase). Implications for mitochondrial cytochrome P-450 systems.";
 Eur. J. Biochem. 169:449-455(1987).
 [6]
 RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 33-492.
 RP TISSUE-Adrenal gland;
 RX MEDLINE-99299392; PubMed-10369776;
 RA Ziegler G.A., Vonnheim C., Hanukoglu I., Schulz G.E.;
 RT "The structure of adrenodoxin reductase of mitochondrial P450 systems:
 electron transfer for steroid biosynthesis.";
 J. Mol. Biol. 289:981-990(1999).
 CC -1- FUNCTION. SERVES AS THE FIRST ELECTRON TRANSFER PROTEIN IN ALL THE
 CC MITOCHONDRIAL P450 SYSTEMS. INCLUDING CHOLESTEROL SIDE CHAIN
 CC CLEAVAGE IN ALL STEROIDOGENIC TISSUES, STEROID 11-BETA
 CC HYDROXYLATION IN THE ADRENAL CORTEX, 25-OH-VITAMIN D3-24
 CC HYDROXYLATION IN THE KIDNEY, AND STEROL C-27 HYDROXYLATION IN THE
 CC LIVER.
 CC -1- CATALYTIC ACTIVITY. REDUCED ADRENODOXIN + NADP(+) - OXIDIZED
 CC ADRENODOXIN + NADPH.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: CHOLESTEROL SIDE-CHAIN-CLEAVAGE SYSTEM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS. A SHORT FORM (SHOWN HERE) AND A
 CC LONG FORM. ARE PRODUCED BY ALTERNATIVE SPLICING. THE LONG FORM
 CC REPRESENTS 10-20% OF ALL ADRENODOXIN REDUCTASE MRNA. AND SEEMS TO
 CC BE INACTIVE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC The European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC DR EMBL: M17029; AAA30362.1; -;
 CC DR EMBL: D00211; BAA00150.1; -;
 CC DR EMBL: X13736; CAA32002.1; -;
 CC DR PIR: A29604; A29604.
 CC DR PIR: JS0390; JS0390.
 CC DR PIR: S03558; S03558.
 CC DR PIR: J70751; J70751.
 CC DR PDB: 1CJC; 12-APR-99.
 CC DR PDB: 1E1L; 02-JUN-00.
 CC DR InterPro: IPR000759; -;
 CC DR PRINTS: PR00419; ADXRDPASE.
 CC DR Election transport; Oxidoreductase: Flavoprotein; NADP: FAD;
 CC Mitochondrion; Transl. peptide; Alternative splicing; 3D-structure.
 CC KW TRANSPORT
 CC FT CHAIN 1 32
 CC FT VARSPIC 33 492
 CC FT CONFLICT 204 204
 CC FT CONFLICT 77 77
 CC FT CONFLICT 81 94
 CC FT CONFLICT 124 128
 CC FT CONFLICT 268 268
 CC FT CONFLICT 317 318
 CC FT CONFLICT 323 323
 CC FT CONFLICT 341 352
 CC SEQUENCE 492 AA: 54338 MW: E66F6F5D18F5131 CR664;
 ODAYH -> RYRLT (IN REF. 3).
 K -> R (IN REF. 3).
 P -> RL (IN REF. 3).
 RAAGRIAVLT -> ARRSAMQSP (IN REF. 3).
 TRAVPTGVDEL -> HPSAHMGCGP (IN REF. 3).
 E66F6F5D18F5131 CR664;

Query Match	53.8%	Score 49;	DB 1;	Length 492;
Best Local Similarity	83.3%	Pred. No. 5.2;		
Matches	5;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0;

```

OY      5 WPMWPM 10
        1 1111
Db      6 WRMWPM 11

RESULT  4
YD55_MYCTU
ID      YD55_MYCTU      STANDARD:      PRT:      715 AA.
AC      011025;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      HYPOTHETICAL 78.2 KDA PROTEIN RV1355C.
GN      RV1355C OR MTCY02B10.19C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxId=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN:H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eiglmeier K., Gass S., Barry C.E. III, Tekala F.,
RA      Badcock R., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squires J., Sulston J.E.,
RA      Taylor K., Whitehead S., Barrall B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."
RL      Nature 393:537-544(1998).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Z75555; CAAG9988.1; -.
DR      Tuberculist: RV1355C; -.
DR      InterPro: IPR000594; -.
DR      Pfam: PF00899; Thif_family: 1.
DR      Hypothetical protein.
KW      Hypothetical protein.
SQ      SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;

Query Match      53.8%; Score 49; DB 1; Length 715;
Best Local Similarity 60.0%; Pred. No. 7.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0.

OY      3 KKMPPMWR 12
        1:1::1111
Db      64 KRMAYPMWR 73

RESULT  5
Y945_MYCTU
ID      Y945_MYCTU      STANDARD:      PRT:      253 AA.
AC      P71564;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      PUTATIVE OXIDOREDUCTASE RV0945 (EC 1.-.-.-).
GN      RV0945 OR MTCY10D7.29C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxId=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN:H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eiglmeier K., Gass S., Barry C.E. III, Tekala F.,
RA      Badcock R., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squires J., Sulston J.E.,
RA      Taylor K., Whitehead S., Barrall B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."
RL      Nature 393:537-544(1998).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: Z75555; CAAG9988.1; -.
DR      Tuberculist: RV1355C; -.
DR      InterPro: IPR000594; -.
DR      Pfam: PF00899; Thif_family: 1.
DR      Hypothetical protein.
KW      Hypothetical protein.
SQ      SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;

```

```

Query Match          53.8%; Score 49; DB 1; Length 715;
Best Local Similarity 60.0%; Pred. No. 7.4;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY      3 KKPMWPMWR 12
        ||:|||||
DB      64 KRMAYPMWR 73

RESULT 5
Y945_MYCFU
ID      Y945_MYCTU      STANDARD:      PRT:      253 AA.
AC      P71564:
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      30-MAY-2000 (Rel. 35, Last annotation update)
DE      PUTATIVE OXIDOREDUCTASE RV0945 (EC 1.-.-.-).
GN      RV0945 OR MTCY10D7.29C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxId=1773;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV:
RX MEDLINE-98295987: PubMed-9634230:
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagsen K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RL complete genome sequence."
CC Nature 393:537-544(1998).
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: 279700; CAB02005.1; -
CC DR TubercuList: RV0945; -
CC DR InterPro: IPR002198; -
CC DR Pfam: PF00106; adh_short; 1.
CC DR PROSITE: PS00061; ADH_SHORT; 1.
CC DR Hypothetical protein; Oxidoreductase.
CC FT ACT_SITE 159 159 BY SIMILARITY.
CC FT SEQUENCE 253 AA; 27138 MW; BAD937208842DA12 CRC64;
CC -----
Query Match 51.6%; Score 47; DB 1; Length 253;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 PWWPW 10
DB 230 PWWPW 234
VG2_IBVD2
ID VG2_IBVD2 STANDARD; PRT; 1154 AA.
AC P12722; Q66176; Q66177;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)
DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].
GN S.
OS Avian infectious bronchitis virus (strain D274) (IBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11124;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-89386000: PubMed-2550899;
RA Jordt B.J.A.M., Kremers D.A.W.M., Kusters H.G., van der Zeijst B.A.M.;
RT "Nucleotide sequence of the gene coding for the peplomer protein (-
RT spike protein) of infectious bronchitis virus, strain D274."
RT Nucleic Acids Res. 17:6726-6726(1989).
CC -1- FUNCTION: THE PELOMER PROTEIN MEDIATES THE BINDING OF VIRIONS
CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC DR EMBL: X15832; CA33837.1; -
CC DR PIR: A34300; VGHIB.
CC DR InterPro: IPR002551; -
CC DR InterPro: IPR002552; -
CC DR Pfam: PF01600; Corona_S1; 1.
CC DR Pfam: PF01601; Corona_S2; 1.
CC DR Glycoprotein; Envelope protein; Transmembrane; Signal.
CC FT SIGNAL 1 18
CC FT CHAIN 19 1154 E2 GLYCOPROTEIN.
CC FT CHAIN 19 538 SPIKE PROTEIN S1.
CC FT CHAIN 539 1154 SPIKE PROTEIN S2.
CC FT DOMAIN 1121 1138 CYS-RICH.
CC FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 592 592 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 948 948 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 961 961 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 1052 1052 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT CARBOHYD 1075 1075 N-LINKED (GLCNAC. . .) (POTENTIAL)
CC FT SEQUENCE 1154 AA; 127502 MW; D79F37AF891A37F CRC64;
CC -----
Query Match 51.1%; Score 46.5; DB 1; Length 1154;
Best Local Similarity 61.5%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;
QY 1 IIK---KMPWPM 10
DB 1086 IIKTYIKMPWYVW 1098
VG2_IBVD
ID VG2_IBVD STANDARD; PRT; 1162 AA.
AC P11223; P05134;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)
DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].
GN S.
OS Avian infectious bronchitis virus (strain Beaudette) (IBV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11122;
RN [1]

```

SEQUENCE FROM N.A.
 MEDLINE=85159540; PubMed=2984314;
 Bluns M.M., Boursnell M.E.G., Cavanagh D., Peppind D.J.C.,
 Brown T.D.K.;
 "Cloning and sequencing of the gene encoding the spike protein of the
 coronavirus IBV";
 J. Gen. Virol. 66:719-726(1985).
 [2]
 SEQUENCE FROM N.A.
 MEDLINE=87085499; PubMed=3025348;
 Bluns M.M., Boursnell M.E.G., Tomley F.M., Brown D.K.;
 "Comparison of the spike precursor sequences of coronavirus IBV
 strains M41 and 6/82 with that of IBV Beaudette";
 J. Gen. Virol. 67:2825-2831(1986).
 -1- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC
 DR EMBL; M91519; AAA70235.1; -
 DR EMBL; X02342; CAA26201.1; -
 DR InterPro: IPR002551; -
 DR InterPro: IPR002552; -
 DR Pfam; PF01600; Corona_S1; 1.
 DR Pfam; PF01601; Corona_S2; 1.
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1162 E2 GLYCOPROTEIN.
 FT CHAIN 19 537 SPIKE PROTEIN S1.
 FT DOMAIN 1120 1137 SPIKE PROTEIN S2.
 FT CYS-RICH 51 51
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1162 1162 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1162 AA; 128046 MW; 0BAAD58113CB8BDS CRR64;

OY 1 ILK---KMPMPW 10

Query Match 51.1%; Score 46.5; DB 1; Length 1162;
 Best Local Similarity 61.5%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

DB 1085 ILKTYIKMPWVW 1097
 ||| |||||
 RESULT 8
 VGL2_IBVK STANDARD; PRT; 1162 AA.
 ID VGL2_IBVK
 AC P12650;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN)
 DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].
 GN S.
 OS Avian infectious bronchitis virus (strain KB8523) (IBV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 OC NCBI_TaxID=11126;
 RX MEDLINE=88306251; PubMed=2841803;
 RA Saitou S., Sato S., Okabe T., Nakai M., Sasaki N.;
 "Cloning and sequencing of genes encoding structural proteins of
 avian infectious bronchitis virus";
 J. Virology 165:589-595(1988).
 FT -1- FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 FT TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.
 FT -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL; M21515; AAA66578.1; -
 DR PIR; B29249; VGIAHK.
 DR InterPro: IPR002551; -
 DR InterPro: IPR002552; -
 DR Pfam; PF01600; Corona_S1; 1.
 DR Pfam; PF01601; Corona_S2; 1.
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 1162 E2 GLYCOPROTEIN.
 FT CHAIN 19 537 SPIKE PROTEIN S1.
 FT DOMAIN 1120 1137 SPIKE PROTEIN S2.
 FT CYS-RICH 51 51
 FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1058 1058 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1162 AA: 128537 MW: 229903683597EABF CRC64:

Query Match 51.1%; Score 46.5; DB 1; Length 1162;
 Best Local Similarity 61.5%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 ILK---KMPMPW 10
 1085 ILKTYIKMPWYV 1097

RESULT 9
 VGL2_IBV6 STANDARD; PRT; 1162 AA.

AC P12651;
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 13-JUL-1999 (Rel. 38, Last annotation update)
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)
 DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].
 GN S.
 OS Avian infectious bronchitis virus (strain M41) (IBV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.

CC NCBI_TaxID=11127;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE-87021475; PubMed-2429473;
 CC Niesters H.G.M., Ienstra J.A., Spaan W.J.M., Zijgerveld A.J.,
 CC Blumink-Pluym N.M.C., Hong F., van Scharrenburg G.J.M.,
 CC Horzinek M.C., van der Zeijst B.A.M.;
 CC "The peplomer protein sequence of the M41 strain of coronavirus IBV
 CC and its comparison with Beaudette strains."
 CC Virus Res. 5:253-263(1986).
 CC -1- FUNCTION: THE PELOMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M21883; AAA6575.1; -
 CC EMBL: A24863; CAA01736.1; -
 CC PIR: S07421; S07421.
 CC InterPro: IPR002551; -
 CC DR pfam: PF01600; Corona_S1; 1.
 CC DR pfam: PF01601; Corona_S2; 1.
 CC Glycoprotein; Envelope protein; Transmembrane; Signal.
 CC SIGNAL
 CC FT CHAIN 1 18
 CC FT CHAIN 19 1162 E2 GLYCOPROTEIN.
 CC FT CHAIN 538 1162 SPIKE PROTEIN S1.
 CC FT DOMAIN 1120 1137 SPIKE PROTEIN S2.
 CC FT CARBOHYD 51 51 CYS-RICH.
 CC FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 247 247 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 425 425 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 513 513 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 530 530 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 579 579 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 947 947 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 979 979 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1014 1014 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1038 1038 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1051 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1162 AA: 128077 MW: 3C9CC70938492DDA CRC64:

Query Match 51.1%; Score 46.5; DB 1; Length 1162;
 Best Local Similarity 61.5%; Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 1 ILK---KMPMPW 10
 1085 ILKTYIKMPWYV 1097

RESULT 10
 VGL2_IBV6 STANDARD; PRT; 1163 AA.

AC P05135;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLOMER PROTEIN)
 DE [CONTAINS: SPIKE PROTEIN S1; SPIKE PROTEIN S2].
 GN S.
 OS Avian infectious bronchitis virus (strain 6/82) (IBV).
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 CC NCBI_TaxID=11121;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE-87085499; PubMed-3025348;
 CC Blum M.M., Bousnell M.E.G., Tomley F.M., Brown D.K.;
 CC "Comparison of the spike precursor sequences of coronavirus IBV
 CC strains M41 and 6/82 with that of IBV Beaudette."
 CC J. Gen. Virol. 67:2825-2831(1986).
 CC -1- FUNCTION: THE PELOMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 CC TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X04723; CAA28432.1; -
 CC InterPro: IPR002551; -
 CC DR pfam: PF01600; Corona_S1; 1.
 CC DR pfam: PF01601; Corona_S2; 1.
 CC Glycoprotein; Transmembrane; Signal.
 CC SIGNAL
 CC FT CHAIN 1 18
 CC FT CHAIN 19 1163 E2 GLYCOPROTEIN.
 CC FT CHAIN 539 1163 SPIKE PROTEIN S1.
 CC FT CHAIN 539 1163 SPIKE PROTEIN S2.

FT DOMAIN 1121 1138 CYS-RICH
 FT CARBOHYD 23 23 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 164 164 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 213 213 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 426 426 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 514 514 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 580 580 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 592 592 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 677 677 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 948 948 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 961 961 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 1052 1052 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 1075 1075 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 1163 AA: 128684 MW: 8FE34ACFE2995478C CRC64:
 SO SEQUENCE

Query Match 51.1% Score 46.5; DB 1; Length 1163;
 Best Local Similarity 61.5% Pred. No. 25;
 Matches 8; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

OY 1 ILK---KMPWMPW 10
 DB 1086 ILKTYIKMPYVW 1098

RESULT 11
 YAO5-SCHPO STANDARD; PRT; 196 AA.
 ID YAO5-SCHPO
 AC 009677;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOHETICAL 22.1 KDA PROTEIN CSH10.05C IN CHROMOSOME I.
 GN SPAC5H10.05C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetes;
 OC Schizosaccharomycetes;
 OC NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: STRONG, TO BACTERIAL MODULATOR OF DRUG ACTIVITY B
 CC (MDAB).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@isb-sib.ch).

CC EMBL: 249811; CAAB9955.1;
 DR Hypothetical protein.
 KW SEQUENCE 196 AA: 22104 MW: 436764DA9E26074C CRC64:
 SO

Query Match 50.5% Score 46; DB 1; Length 196;
 Best Local Similarity 47.1% Pred. No. 5.7;
 Matches 8; Conservative 3; Mismatches 2; Indels 4; Gaps 2;

OY 1 ILKWP-MW---PWRRK 13
 DB 62 ILYQPMGMMGTWKLK 78

RESULT 12
 MML6-MYCTU STANDARD; PRT; 397 AA.
 ID MML6-MYCTU
 AC 010773;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN MML6.
 GN MML6 OR RV1557 OR MTCY48.08C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Filicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OC NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RC MEDLINE=963295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Barry C.E. III, Tekaia F.,
 RA Gordon S.V., Eigmeier K., Gas S., Chillingworth T., Connor R.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jorgels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter K., Seeger K., Skellon S., Squares S., Squires R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RT Nature 393:537-544(1998).
 RL Nature 393:537-544(1998).
 CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1 SIMILARITY: BELONGS TO THE MML FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@isb-sib.ch).

CC EMBL: 274020; CA98334.1;
 DR Tuberculin; RV1557;
 DR Hypothetical protein; Transmembrane.
 KW TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 190 210 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 FT SEQUENCE 397 AA: 42421 MW: 678DC68E24472BF4 CRC64:
 SO

Query Match 49.5% Score 45; DB 1; Length 397;
 Best Local Similarity 54.5% Pred. No. 15;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ILKMPWMPW 11
 DB 348 ILGRWMPWMPOR 358

RESULT 13
 TRPE_PSESS STANDARD; PRT: 505 AA.
 ID TRPE_PSESS
 AC P21689; 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 01-MAY-1991 (Rel. 18, Last annotation update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE ANTHRANILATE SYNTHASE COMPONENT I (EC 4.1.3.27).
 GN TRPE.
 OS Pseudomonas syringae (pv. savastanoi).
 CC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 CC NCBI_TaxID=29438;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91100331; Pubmed-1987141;
 RA da Costa E., Silva O., Kosuge T.;
 RT "Molecular characterization and expression analysis of the
 anthranilate synthase gene of Pseudomonas syringae subsp.
 savastanoi." 173:463-471(1991).
 RT J. Bacteriol. 173:463-471(1991).
 CC -1- CATALYTIC ACTIVITY: CHORISMATE + L-GLUTAMINE - ANTHRANILATE +
 PYRUVATE + L-GLUTAMATE.
 CC -1- PATHWAY: FIRST STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC -1- SUBUNIT: TETRAMER OF TWO COMPONENTS I AND TWO COMPONENTS II (BY
 SIMILARITY).
 CC -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE
 USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES
 GLUTAMINE AMIDOTRANSFERASE ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M55911; AAA26016.1; -
 DR PIR: A39128; A39128.
 DR InterPro: IPR000350; -
 DR Pfam: PF00425; chorismate_bind; 1.
 DR PRINTS: PR00095; ANTSNTHASEI.
 DR Trypophan biosynthesis; Lyase.
 KW Trypophan biosynthesis; Lyase.
 SQ SEQUENCE 505 AA; 56084 MW; A38E81931331F6BB CRC64;

Query Match 49.5%; Score 45; DB 1; Length 505;
 Best Local Similarity 71.4%; Pred. No. 19;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

RP SEQUENCE FROM N.A.
 RA Cedberg H., Hohmann S., Schaaff-Gerstenschlaeger I., Huse K.,
 RA Zimmermann F.K.;
 RA Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93070619; Pubmed-1332309;
 RA Cardone M.L.A., Panzeri L., Falconi M.M., Carcano C., Plevani P.,
 RA Lucchini G.;
 RT "Nucleotide sequence of 9.2 kb left of CRY1 on yeast chromosome III
 from strain AB972: evidence for a Ty insertion and functional
 analysis of open reading frame YCR28.";
 RT Yeast 8:805-812(1992).
 RN [3]
 RP SIMILARITY TO DAL5 FAMILY
 RX MEDLINE-94147996; Pubmed-8313894;
 RA Koonin E.V., Bork P., Sander C.;
 RT "Yeast chromosome III: new gene functions.";
 RN [4]
 RP CHARACTERIZATION.
 RX MEDLINE-96367594; Pubmed-8771708;
 RA Marchreau C., Joets J., Pousset D., Guilloton M., Karst F.;
 RT "FEN2: a gene implicated in the catabolite repression-mediated
 regulation of ergosterol biosynthesis in yeast.";
 RN [5]
 RT Yeast 12:531-539(1996).
 CC -1- FUNCTION: INVOLVED IN THE CATABOLITE REPRESSION-MEDIATED
 REGULATION OF ERGOSTEROL BIOSYNTHESIS AND IN FENPROPIOMORPH
 RESISTANCE.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE ALANTOATE PERMEASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X59720; CAA42320.1; -
 DR PIR: S19439; S19439.
 DR PIR: S25336; S25336.
 DR SGD: S0000623; FEN2.
 DR Transmembrane; Transport.
 KW Transmembrane; Transport.
 FT TRANSMEM 28 48
 FT TRANSMEM 80 100
 FT TRANSMEM 103 123
 FT TRANSMEM 133 153
 FT TRANSMEM 165 185
 FT TRANSMEM 199 219
 FT TRANSMEM 272 292
 FT TRANSMEM 313 333
 FT TRANSMEM 363 383
 FT TRANSMEM 393 413
 FT TRANSMEM 422 442
 FT TRANSMEM 455 475
 FT TRANSMEM 485 505
 FT CONFLICT 104 104
 FT SEQUENCE 512 AA; 58256 MW; 361942E74C62B5B4 CRC64;

Query Match 49.5%; Score 45; DB 1; Length 512;
 Best Local Similarity 62.5%; Pred. No. 19;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 IIRKPMW 8
 DB 268 VIKRHW 275

RESULT 15
 MML5_MYCTU


```

ID      MML5_MYCTU      STANDARD:      PRT:      964 AA.
AC      053784:
DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DT      30-MAY-2000 (Rel. 39, Last annotation update)
DE      PUTATIVE MEMBRANE PROTEIN MML5.
GN      MML5 OR RV0676C OR MT040.04C.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J.E.,
RA      Taylor K., Whitehead S., Barrell B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."
RL      Nature 393:537-544(1998).
CC      -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC      -1- SIMILARITY: BELONGS TO THE MML FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL: AL021943; CAA17459.1; -
DR      Tuberculist; RV0676c; -
DR      InterPro: IPR001036; -
DR      PRINTS: PRO0702; ACRIFLAVINRP.
KW      Hypothetical protein; Transmembrane.
FT      TRANSMEM 31 51 POTENTIAL.
FT      TRANSMEM 203 223 POTENTIAL.
FT      TRANSMEM 230 250 POTENTIAL.
FT      TRANSMEM 255 275 POTENTIAL.
FT      TRANSMEM 302 322 POTENTIAL.
FT      TRANSMEM 340 360 POTENTIAL.
FT      TRANSMEM 389 409 POTENTIAL.
FT      TRANSMEM 745 765 POTENTIAL.
FT      TRANSMEM 774 794 POTENTIAL.
FT      TRANSMEM 803 823 POTENTIAL.
FT      TRANSMEM 826 846 POTENTIAL.
FT      TRANSMEM 880 900 POTENTIAL.
FT      TRANSMEM 901 921 POTENTIAL.
SQ      SEQUENCE 964 AA; 104784 MW; B7C945940A1176BD CRC64;
Query Match 49.5%; Score 45; DB 1; Length 964;
Best Local Similarity 66.7%; Pred. NO. 34;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY      1 ILKKPMP 9
DB      932 LIGKFWMP 940

```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:33:38 ; Search time 33.4 Seconds
(without alignments)
51.496 Million cell updates/sec

Title: US-09-444-281-35
Perfect score: 91
Sequence: 1 ILKKPMWPMRRK 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.16:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_unclassified:*
13: sp_vertebrate:*
14: sp_virus:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	62.6	723	14 Q9DUC4	Q9DUC4 tt virus. O
2	54	59.3	1383	14 Q84712	Q84712 porcine epi
3	52	57.1	746	14 Q9JH31	Q9JH31 tt virus. O
4	52	57.1	1018	1 Q9HXX3	Q9HXX3 thermoplas
5	51	56.0	299	4 Q9Y4N1	Q9Y4N1 homo sapien
6	49	53.8	148	5 Q26590	Q26590 schistosoma
7	49	53.8	298	1 Q9Y806	Q9Y806 aeropyrum p
8	49	53.8	467	5 Q19573	Q19573 caenorhabdi
9	49	53.8	528	5 Q26589	Q26589 schistosoma
10	49	53.8	528	5 Q9TY57	Q9TY57 schistosoma
11	49	53.8	735	14 Q9DUC9	Q9DUC9 tt virus. O
12	49	53.8	802	5 Q96398	Q96398 schistosoma
13	49	53.8	1245	3 Q9Y7V5	Q9Y7V5 trichoderma
14	49	53.8	1940	5 Q02456	Q02456 schistosoma
15	48	52.7	49	14 Q9D780	Q9D780 tt virus. O
16	48	52.7	111	5 Q18753	Q18753 caenorhabdi
17	48	52.7	428	11 Q9JMG0	Q9JMG0 mus musculi
18	48	52.7	748	14 Q9D781	Q9D781 tt virus. O
19	47.5	52.2	114	2 Q9X8C2	Q9X8C2 streptomyce

20	47	51.6	165	10 Q9SNM3	Q9SNM3 Oryza sativ
21	47	51.6	504	2 P96143	P96143 thermocactin
22	46.5	51.1	352	2 P73417	P73417 synchocyst
23	46.5	51.1	620	14 Q9YH07	Q9YH07 avian infec
24	46.5	51.1	621	14 Q9YH07	Q9YH07 avian infec
25	46.5	51.1	621	14 Q9YH14	Q9YH14 avian infec
26	46.5	51.1	621	14 Q9YH13	Q9YH13 avian infec
27	46.5	51.1	621	14 Q9YH11	Q9YH11 avian infec
28	46.5	51.1	621	14 Q9YH10	Q9YH10 avian infec
29	46.5	51.1	621	14 Q9YH15	Q9YH15 avian infec
30	46.5	51.1	621	14 Q9YH12	Q9YH12 avian infec
31	46.5	51.1	621	14 Q9YH09	Q9YH09 avian infec
32	46.5	51.1	621	14 Q9YH08	Q9YH08 avian infec
33	46.5	51.1	625	14 Q9YH06	Q9YH06 avian infec
34	46.5	51.1	630	14 Q9YH17	Q9YH17 avian infec
35	46.5	51.1	630	14 Q9YH19	Q9YH19 avian infec
36	46.5	51.1	630	14 Q9YH20	Q9YH20 avian infec
37	46.5	51.1	630	14 Q9YH21	Q9YH21 avian infec
38	46.5	51.1	630	14 Q9YH22	Q9YH22 avian infec
39	46.5	51.1	630	14 Q9YH23	Q9YH23 avian infec
40	46.5	51.1	630	14 Q9YH24	Q9YH24 avian infec
41	46.5	51.1	630	14 Q9YH25	Q9YH25 avian infec
42	46.5	51.1	630	14 Q9YH26	Q9YH26 avian infec
43	46.5	51.1	630	14 Q9YH27	Q9YH27 avian infec
44	46.5	51.1	630	14 Q9YH28	Q9YH28 avian infec
45	46.5	51.1	630	14 Q9YH29	Q9YH29 avian infec

ALIGNMENTS

RESULT 1
ID Q9DUC4 PRELIMINARY; PRT; 723 AA.
AC Q9DUC4;
DC 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE ORF1.
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_Taxid=68887;
RN [1]
RP SEQUENCE FROM N.A.
RA Okamoto H.;
RC STRAIN=MF-TTV9;
RA Okamoto H.;
RN Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MF-TTV9;
RX PubMed=11080484;
RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
RA Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
RT "Species-specific TT viruses in humans and nonhuman primates and their
RT phylogenetic relatedness.";
RL Virology 277:368-378(2000).
DR EMBL; AB041959; BAB19313.1;
SQ SEQUENCE 723 AA; 85393 MW; 232D003098766344 CRC64;

Query Match 62.6%; Score 57; DB 14; Length 723;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 PMWPMR 12
Db 2 PMWPMR 8
RESULT 2
ID Q84712 PRELIMINARY; PRT; 1383 AA.
AC Q84712;

DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE SPIKE PROTEIN.
 OS Porcine epidemic diarrhoea virus.
 OS Virus; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronaviruses.
 OC NCBI_TaxID=28295;
 OX NCBI_TaxID=28295;
 RN NCBI_TaxID=28295;
 RP STRAIN-BRI/87;
 RC STRAIN-BRI/87;
 RX MEDLINE-94231173; PubMed-8176382;
 RA Duarte M., Laude H.;
 RT "Sequence of the spike protein of the porcine epidemic diarrhoea virus."
 RT J. Gen. Virol. 75:1195-1200(1994).
 RL J. Gen. Virol. 75:1195-1200(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRI/87;
 RX MEDLINE-93389433; PubMed-8397280;
 RA Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
 RT "Sequence determination of the nucleocapsid protein gene of the porcine epidemic diarrhoea virus confirms that this virus is a coronavirus related to human coronavirus 229E and porcine transmissible gastroenteritis virus."
 RT J. Gen. Virol. 74:1795-1804(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRI/87;
 RX MEDLINE-94120721; PubMed-8291230;
 RA Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M., Laude H.;
 RT "Sequence analysis of the porcine epidemic diarrhoea virus genome reveals the nucleocapsid and spike protein genes reveals a polymorphic ORF."
 RT Virology 198:466-476(1994).
 RL EMBL; Z25483; CAA80971.1; -
 DR InterPro: IPR002351; -
 DR Pfam: PF01600; Corona_S1; 1.
 DR Pfam: PF01601; Corona_S2; 1.
 DR CONFLICT 422 422 Y -> N (IN REF. 1).
 FT SEQUENCE 1383 AA; 151404 MW; 741C84D5D3BDC4D CRC64;
 SO SEQUENCE

Query Match 59.3%; Score 54; DB 14; Length 1383;
 Best Local Similarity 85.7%; Pred. NO. 17;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 KMPWMP 10
 DB 1322 KMPWMP 1328

RESULT 3
 ID Q9JH31 PRELIMINARY; PRT; 746 AA.
 AC Q9JH31;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE ORF1.
 OS TT virus.
 OS Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OC NCBI_TaxID=68887;
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TJN02;
 RA Okamoto H.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TJN02;

RA Ukita M., Okamoto H., Nishizawa T., Tawara A., Takahashi M., Iizuka H., Miyakawa Y., Mayumi M.;
 RT "The entire nucleotide sequences of two distinct TT virus (TTV) isolates (TJN01 and TJN02) remotely related to the original TTV isolates."
 RT J. Virol. 74:1795-1804(1993).
 RL Arch. Virol. 0:0-0(2000).
 DR EMBL; AB028669; BAA94878.1; -
 SO SEQUENCE 746 AA; 88561 MW; E0B22953AE764E3E CRC64;

Query Match 57.1%; Score 52; DB 14; Length 746;
 Best Local Similarity 66.7%; Pred. NO. 18;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 WPMWPRK 13
 DB 3 WPMWPRK 11

RESULT 4
 ID Q9HKX3 PRELIMINARY; PRT; 1018 AA.
 AC Q9HKX3;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE CONSERVED HYPOTHETICAL MEMBRANE PROTEIN.
 GN TA0470.
 OS Thermoplasma acidophilum.
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;
 OC Thermoplasma.
 OX NCBI_TaxID=2303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DSM 1728;
 RX MEDLINE-20479972; PubMed-11029001;
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C., Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumelster W.;
 RT "The genome sequence of the thermophilic scavenger Thermoplasma acidophilum."
 RT Nature 407:508-513(2000).
 RL EMBL; AL445064; CAC11612.1; -
 DR InterPro: IPR000731; -
 SO SEQUENCE 1018 AA; 112323 MW; 83EE84D3C74B852 CRC64;

Query Match 57.1%; Score 52; DB 1; Length 1018;
 Best Local Similarity 66.7%; Pred. NO. 24;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILKMPWP 9
 DB 1004 ILKMPWP 1012

RESULT 5
 ID Q9Y4N1 PRELIMINARY; PRT; 299 AA.
 AC Q9Y4N1;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE HYPOTHETICAL 34.0 KDA PROTEIN (FRAGMENT).
 GN DKFZ434C192.
 OS Homo sapiens (human).
 OS Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RA Ansoorge W., Winkner U., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL: AL096753; CAB46428.2; -
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 299 AA; 34032 MW; 6B8DB606A88239A CRC64;

Query Match 56.0%; Score 51; DB 4; Length 299;
 Best Local Similarity 85.7%; Pred. No. 11;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 6 PMPMPMR 12
 |||||
 Db 37 PMPMPMR 43

RESULT 6
 026590 PRELIMINARY; PRT; 148 AA.

AC 026590; (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE MYOSIN HEAVY CHAIN (FRAGMENT).
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigoidida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OX NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Schmitz J., Symons P., Dargatz H., Kunz W.;
 RL Submitted (FEB-1992) to the EMBL/GenBank/DBJ databases.
 DR EMBL: M81338; AAA29906.1;
 KW Myosin.
 FT NON_TER 1 148
 FT NON_TER 148
 SQ SEQUENCE 148 AA; 17923 MW; C7EDA5A0BBE14DDA CRC64;

Query Match 53.8%; Score 49; DB 5; Length 148;
 Best Local Similarity 62.5%; Pred. No. 10;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILKMPMW 8
 |||||
 Db 30 VLKMPMW 37

RESULT 7
 09Y806 PRELIMINARY; PRT; 298 AA.
 ID 09Y806; (TREMBlrel. 12, Created)
 AC 09Y806; (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE HYPOTHETICAL 33.7 KDA PROTEIN APE2577.
 GN APE2577.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,
 RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankal A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
 RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 crenarchaeon, Aeropyrum pernix K1.";

RL DNA Res. 6:83-101(1999).
 DR EMBL: AP000064; BAA81594.1; -
 DR InterPro: IPR002787; -
 DR Pfam: PF01932; DUF85; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 298 AA; 33666 MW; FCB9C6EC93FF231 CRC64;

Query Match 53.8%; Score 49; DB 1; Length 298;
 Best Local Similarity 60.0%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 LKMPMPMR 11
 |||||
 Db 102 IKETPMPMR 111

RESULT 8
 019573 PRELIMINARY; PRT; 467 AA.

AC 019573; (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE STICKER TO 9 AMINO ACID REPEATS IN GALACTOSE SPECIFIC LECTINS.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94150718; PubMed=7906398;
 RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
 RA Bonfield J., Burton J., Connell M., Cosey T., Cooper J., Coulson A.,
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 RA Gardner A., Green P., Hawkins T., Hillier L., Jiet M., Johnston L.,
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
 RA Smailson N., Smith A., Sonhammer E., Staden R., Sulston J.,
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Woldman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.";
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Favello T.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U39855; AAA81082.1; -
 KW Lectin.
 SQ SEQUENCE 467 AA; 53169 MW; 7D9BBAB61830431B CRC64;

Query Match 53.8%; Score 49; DB 5; Length 467;
 Best Local Similarity 83.3%; Pred. No. 29;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 WPMWPMW 10
 |||||
 Db 201 WPMWPMW 206

RESULT 9
 026589 PRELIMINARY; PRT; 528 AA.
 ID 026589; (TREMBlrel. 01, Created)
 AC 026589; (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE MYOSIN II HEAVY CHAIN (FRAGMENT).
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OC NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PUERTO RICAN;
 RA Amory L.M.;
 RL Submitted (APR-1992) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PUERTO RICAN;
 RA MEDLINE-9305536; PubMed-1431131;
 RX Solis L.M.; Masterson C.P.; Tom T.D.; McNally M.T.; Lowell G.H.;
 RA Strand M.;
 RT "Induction of protective immunity in mice using a 62-kDa recombinant
 fragment of a Schistosoma mansoni surface antigen."
 RT J. Immunol. 149:3612-3620(1992).
 RL EMBL: X65591; CAA46548.1; -
 DR HSP; P08799; 1MMD.
 DR InterPro: IPR000048; -
 DR InterPro: IPR000533; -
 DR InterPro: IPR001609; -
 DR InterPro: IPR001637; -
 DR InterPro: IPR002928; -
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00194; TROPOMYOSIN.
 DR PROSITE: PS00182; GLNA_ADENYLATION; UNKNOWN_1.
 DR SMART; SM00015; IQ; 1.
 DR MYOSIN.
 DR NON_TER
 FT SEQUENCE 528 AA; 61622 MW; AF075D13EB249B4C CRC64;
 SQ

Query Match 53.8%; Score 49; DB 5; Length 528;
 Best Local Similarity 62.5%; Pred. No. 32;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ILKKPMM 8
 DB 106 VLKRPMM 113

RESULT 10
 Q9TY57 PRELIMINARY; PRT; 528 AA.
 AC Q9TY57.
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE MYOSIN HEAVY CHAIN (FRAGMENT).
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OC NCBI_TaxID=6182;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CHINESE MAINLAND STRAIN;
 RX MEDLINE-9914454; PubMed-990643;
 RX Zhang Y.B.; Taylor M.G.; Bickle Q.D.;
 RT "Schistosoma japonicum myosin: cloning, expression and vaccination
 studies with the homologue of the S.mansoni myosin fragment Irv-5."
 RT Parasite Immunol. 20:583-594(1998).
 RL EMBL: U55133; AAC82221.1; -
 DR HSP; P08799; 1MMD.
 DR InterPro: IPR000048; -

DR InterPro: IPR000533; -
 DR InterPro: IPR001609; -
 DR InterPro: IPR002928; -
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00194; TROPOMYOSIN.
 DR SMART; SM00015; IQ; 1.
 DR NON_TER
 FT SEQUENCE 528 AA; 61406 MW; C54A31F540F5EE05 CRC64;
 SQ

Query Match 53.8%; Score 49; DB 5; Length 528;
 Best Local Similarity 62.5%; Pred. No. 32;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ILKKPMM 8
 DB 106 VLKRPMM 113

RESULT 11
 Q9DUC9 PRELIMINARY; PRT; 735 AA.
 AC Q9DUC9.
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OC NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PT-TTV6;
 RA Okamoto H.;
 RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PT-TTV6;
 RX PubMed-11080484;
 RX Okamoto H.; Nishizawa T.; Tawara A.; Peng Y.; Takahashi M.;
 RA Kishimoto J.; Tanaka T.; Miyakawa Y.; Mayumi M.;
 RT "Species-specific TT viruses in humans and nonhuman primates and their
 phylogenetic relatedness."
 RT Virology 277:368-378(2000).
 RL EMBL: AB041957; BAB19308.1; -
 DR SEQUENCE 735 AA; 86132 MW; 9ED818D6B6FA5D3 CRC64;
 SQ

Query Match 53.8%; Score 49; DB 14; Length 735;
 Best Local Similarity 53.8%; Pred. No. 44;
 Matches 7; Conservative 1; Mismatches 1; Indels 4; Gaps 1;
 QY 5 WPMW---WPMRK 13
 DB 3 WPMRRRWRMR 15

RESULT 12
 ID 096398 PRELIMINARY; PRT; 802 AA.
 AC 096398.
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)
 DT 01-MAR-2001 (TREMBLrel. 16, last annotation update)
 DE MYOSIN (FRAGMENT).
 OS Schistosoma japonicum (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabditophora; Neodermata;
 OC Trematoda; Digenea; Strigeidida; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OC NCBI_TaxID=6182;
 QY 096398

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CHINESE MAINLAND;
 RX MEDLINE-99144454; PubMed-9990643;
 RA Zhang Y.B., Taylor M.G., Blackie O.D.;
 RT "Schistosoma japonicum myosin: cloning, expression and vaccination
 studies with the homologue of the S. mansoni myosin fragment Irv-5.";
 RL Parasite Immunol. 20:583-594(1998).
 DR EMBL: AF039187; AAC82332.1; -
 DR HSSP: P08799; 1MD.
 DR InterPro: IPR000048; -
 DR InterPro: IPR000533; -
 DR InterPro: IPR001609; -
 DR InterPro: IPR002928; -
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00194; TROPOMYOSIN.
 DR SMART: SM00015; IQ; 1.
 FT NON_TER 1
 FT NON_TER 802
 SO SEQUENCE 802 AA; 92440 MW; 25F808EBB558FBC6 CRC64;

Query Match 53.8%; Score 49; DB 5; Length 802;
 Best Local Similarity 62.5%; Pred. No. 47;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILKKPMM 8
 Db 146 VLKNPMM 153

RESULT 13
 O9YV5 PRELIMINARY; PRT: 1245 AA.
 AC O9YV5;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE CONIDIOSPORE SURFACE PROTEIN.
 GN Cmpl.
 OS Trichoderma harzianum.
 OC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
 OX NCBI_TaxID=5544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAINATCC 32173;
 RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
 van Montagu M., Herrera Estrella A., Horwitz B.A.;
 RT "Developmental regulation of a gene encoding a multidomain
 conidiospore surface protein of Trichoderma, cpl.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ133651; CAB40845.1; -
 DR SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0CF8 CRC64;

Query Match 53.8%; Score 49; DB 3; Length 1245;
 Best Local Similarity 50.0%; Pred. No. 70;
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 4 KMPWPMRRK 13
 Db 1185 RQWMSWPMRR 1194

RESULT 14
 ID 002456 PRELIMINARY; PRT: 1940 AA.
 AC 002456;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)

DE MYOSIN HEAVY CHAIN.
 GN MYH.
 OS Schistosoma mansoni (Blood fluke).
 OC Eukaryota; Metazoa; Platyhelminthes; Rhabdiorhara; Neodermata;
 OC Trematoda; Digenea; Strigoidae; Schistosomatidae; Schistosomatidae;
 OC Schistosoma.
 OX NCBI_TaxID=6183;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRAZILIAN LE;
 RX MEDLINE-93211444; PubMed-8459827;
 RA Weston D.S., Schmitz J., Kemp M., Kunz M.;
 RT "Cloning and sequencing of a complete myosin heavy chain cDNA from
 Schistosoma mansoni.";
 RL Mol. Biochem. Parasitol. 58:161-164(1993).
 CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
 ACTIVITY THAT IS ACTIVATED BY ACTIN.
 CC -1- SUBUNIT: MUSCLE MYOSIN IS A HEXAMERIC PROTEIN THAT CONSISTS OF 2
 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI LIGHT CHAIN SUBUNITS (MLC)
 AND 2 REGULATORY LIGHT CHAIN SUBUNITS OF THE MYOFIBRILS.
 CC -1- SUBCELLULAR LOCATION: THICK FILAMENTS OF THE MYOFIBRILS.
 CC -1- MISCELLANEOUS: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
 MEROMYOSIN (LMW) AND 1 HEAVY MEROMYOSIN (HMW). IT CAN LATTER BE
 SPLIT FURTHER INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
 SUBFRAGMENT (S2).
 CC -1- MISCELLANEOUS: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE,
 SHOWING CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4
 HEPTAPEPTIDES, CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
 CC -1- SIMILARITY: THE GLOBULAR HEAD AA SEQUENCE SHOWS A HIGH DEGREE OF
 SIMILARITY WITH THE GLOBULAR HEAD AA SEQUENCES OF MUSCLE & NONMUSCLE
 HEAVY CHAINS. BY CONTRAST THE ROD SEQUENCE IS LESS CONSERVED, BUT
 THE PERIODICITIES OF HYDROPHOBIC & CHARGED RESIDUES, WHICH DICTATE
 THE ALPHA-HELICAL COILED-COIL STRUCTURE, ARE CONSERVED.
 DR EMBL: L01634; AAA29905.1; -
 DR HSSP: P08799; 1MD.
 DR InterPro: IPR000048; -
 DR InterPro: IPR000122; -
 DR InterPro: IPR001609; -
 DR InterPro: IPR001637; -
 DR InterPro: IPR002928; -
 DR Pfam: PF00063; myosin_head; 1.
 DR Pfam: PF00612; IQ; 1.
 DR Pfam: PF01576; Myosin_tail; 1.
 DR PRINTS: PR00193; MYOSINHEAVY.
 DR PRODOM: PD000355; -; 1.
 DR PROSITE: PS00182; GLNA_ADENYLATION; UNKNOWN_1.
 DR SMART: SM00015; IQ; 1.
 DR Myosin; Muscle protein; Coiled coil; Actin-binding; ATP-binding;
 KW Heptad repeat pattern.
 FT DOMAIN 1 824
 FT DOMAIN 825 1921
 FT DOMAIN 1922 1940
 FT NP_BIND 172 179
 FT ACT_SITE 682 682
 FT ACT_SITE 692 692
 FT SEQUENCE 1940 AA; 222378 MW; 10FC4EAE208CA365 CRC64;

Query Match 53.8%; Score 49; DB 5; Length 1940;
 Best Local Similarity 62.5%; Pred. No. 1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILKKPMM 8
 Db 809 VLKNPMM 816

RESULT 15
 ID 09DT80 PRELIMINARY; PRT: 49 AA.
 AC 09DT80;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)

DT 01-MAR-2001 (TRENBLREL. 16, last sequence update)
DE 01-MAR-2001 (TRENBLREL. 16, last annotation update)
ORF1 (FRAGMENT).
OS TT virus.
OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
OX NCBI_TaxID=68887;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TY9:
RX MEDLINE-20568739; PubMed-11118348;
RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
RA Sai T., Sugai Y.;
RT "TT virus mRNAs detected in the bone marrow cells from an infected
RT individual.";
RL Biochem. Biophys. Res. Commun. 279:700-707(2000).
DR EMBL, AB050449; BAB19930.1; -.
FT NON_TER 49 49
SQ SEQUENCE 49 AA; 7225 MW; 1DA6F8F1AB69AA43 CRC64;

Query Match 52.7%; Score 48; DB 14; Length 49;
Best Local Similarity 41.2%; Pred. No. 5.2;
Matches 7; Conservative 1; Mismatches 1; Indels 8; Gaps 1;
OY 5 WPV-----MPWRRK 13
1 | | | | |
Db 3 WTWMMQRRRRRMPWRRR 19

Search completed: June 21, 2001, 08:37:43
Job time: 245 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:33:12 ; Search time 30.3 Seconds

(Without alignments)
24.010 Million cell updates/sec

Title: US-09-444-281-36

Sequence: 1 LTRMPMPMPRRK 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID88/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	12	19	AAV24550
2	86	100.0	12	21	AAV94496
3	86	100.0	12	21	AAV91791
4	86	100.0	20	19	AAV24553
5	86	100.0	20	21	AAV91797
6	86	100.0	21	19	AAV24552
7	86	100.0	21	19	AAV24554
8	86	100.0	21	19	AAV66376
9	86	100.0	21	21	AAV91796
10	86	100.0	21	21	AAV91798
11	86	100.0	27	19	AAV66363

12	86	100.0	28	21	AAV91800	Amino acid sequenc
13	83	96.5	12	19	AAV24567	Indolicidin analog
14	83	96.5	12	21	AAV91788	Amino acid sequenc
15	82	95.3	12	19	AAV24594	Indolicidin analog
16	82	95.3	12	19	AAV66364	Indolicidin analog
17	82	95.3	12	21	AAV91817	Amino acid sequenc
18	82	95.3	12	21	AAV91841	Amino acid sequenc
19	81	94.2	12	19	AAV24605	Indolicidin analog
20	81	94.2	12	19	AAV24595	Indolicidin analog
21	81	94.2	12	21	AAV91842	Amino acid sequenc
22	81	94.2	12	21	AAV91852	Amino acid sequenc
23	80	93.0	12	19	AAV24596	Indolicidin analog
24	80	93.0	12	19	AAV24603	Indolicidin analog
25	80	93.0	12	19	AAV24604	Indolicidin analog
26	80	93.0	12	21	AAV91843	Amino acid sequenc
27	80	93.0	12	21	AAV91850	Amino acid sequenc
28	80	93.0	12	21	AAV91851	Amino acid sequenc
29	78	90.7	12	19	AAV24598	Indolicidin analog
30	78	90.7	12	19	AAV24601	Indolicidin analog
31	78	90.7	12	19	AAV66361	Indolicidin analog
32	78	90.7	12	21	AAV91845	Amino acid sequenc
33	78	90.7	12	21	AAV91845	Amino acid sequenc
34	78	90.7	12	21	AAV91848	Amino acid sequenc
35	78	90.7	13	19	AAV24565	Indolicidin analog
36	78	90.7	13	21	AAV91786	Amino acid sequenc
37	77	89.5	12	19	AAV24586	Indolicidin analog
38	77	89.5	12	21	AAV91828	Amino acid sequenc
39	75	87.2	11	19	AAV24569	Indolicidin analog
40	75	87.2	11	21	AAV91790	Amino acid sequenc
41	75	87.2	12	19	AAV24580	Indolicidin analog
42	75	87.2	12	21	AAV91804	Amino acid sequenc
43	75	87.2	13	18	AAV12873	Antimicrobial cati
44	75	87.2	13	18	AAV12895	Antimicrobial cati
45	75	87.2	13	18	AAV12896	Antimicrobial cati

ALIGNMENTS

RESULT 1	
ID	AAV24550
AAV24550	standard; peptide: 12 AA.
AAV24550;	
AC	18-AUG-1999 (first entry)
XX	
XX	
DE	Indolicidin analogue #2.
XX	
KW	Indolicidin; bacterial infection; photo-oxidised solubiliser;
KW	antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;
KW	additive; shampoo; soap; insecticide; herbicide; preservative;
KW	food; technical material.
XX	
OS	Synthetic.
XX	
PN	WO9807745-A2.
PD	26-FEB-1998.
XX	
XX	21-AUG-1997; 97WO-US14779.
PF	
XX	
XX	13-JAN-1997; 97US-0034949.
PR	21-AUG-1996; 96US-0024754.
XX	
PA	(MICR-) MICROLOGIX BIOTECH INC.
XX	
PI	Erfile D, Fraser JR, Krieger TJ, Taylor R, West MH;
XX	WPI; 1998-169090/15.
DR	
XX	New indolicidin analogues with antimicrobial activity and related
PT	nucleic acid - vectors, transformed cells and antibodies, also

Query Match 100.0%; Score 86; DB 21; Length 12;
 Best Local Similarity 100.0%; Pred. No. 3,1e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 ILRMPMPWRRK 12
 |||||
 DB 1 ILRMPMPWRRK 12

RESULT 4

AAV24553

ID AAV24553 standard; peptide; 20 AA.

AC AAV24553;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #5.

Indolicidin; bacterial infection; photo-oxidised solubiliser;
 antimicrobial; antibiotic; antiaerhythmic; surface disinfectant;
 additive; shampoo; soap; insecticide; herbicide; preservative;
 food; technical material.

OS Synthetic.

PN WO9807745-A2.

PD 26-FEB-1998.

PF 21-AUG-1997; 97WO-US14779.

PR 13-JAN-1997; 97US-0034949.

PR 21-AUG-1996; 96US-0024754.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;

DR WPI; 1998-169090/15.

New indolicidin analogues with antimicrobial activity and related
 nucleic acid - vectors, transformed cells and antibodies, also
 conjugates with polyoxyalkylene glycol and fatty acid to reduce
 toxicity, useful therapeutically, as disinfectants etc.

PS Claim 11; Page 88; 129pp; English.

AAV24549 to AAV24615 represent indolicidin analogues of formulae
 (I)-(VIII) containing up to 25 amino acids (aa): RZXZXZB (I), BXZXZXB
 (II), BBBZXZXZB (III), BXZXZBBBn(AA)nMLBBGS (IV), BXZXZXB(AA)nM
 (V), LBBnZXZXZB (VI), LKZXZXZB (VII) and BBZXZXZBBB (VIII).
 Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
 preferably R or K; AA = any aa; n = 0 or 1; In (II), at least 1 Z = V;
 in (VIII) at least 2 X = F or Y. The analogues are used to treat
 infections caused by bacteria (Gram positive or negative, or anaerobic);
 fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 trematodes) or viruses. Typical of very many pathogens that can be
 controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 derived from the analogues may be used similarly; the compounds may
 also be prepared from antibiotics or antiaerhythmic agents. The analogues
 may be used therapeutically or to coat medical devices; also they are
 useful as surface disinfectants, as additives to shampoo or soaps, as
 insecticides or herbicides, or as preservatives for foods and technical
 materials. The analogues are administered by injection, lavage, orally
 or topically, generally at 0.1-50 mg/Kg. These analogues have a broader
 spectrum of activity than indolicidin and modification as compounds
 reduces their toxicity.

SQ Sequence 20 AA;

Query Match 100.0%; Score 86; DB 19; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5,1e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRMPMPWRRK 12
 |||||
 DB 1 ILRMPMPWRRK 12

RESULT 5

AAV91797

ID AAV91797 standard; Peptide; 20 AA.

AC AAV91797;

DT 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11B17CN.

Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
 breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 multidrug resistance.

OS Synthetic.

PN WO965506-A2.

PD 23-DEC-1999.

PF 14-JUN-1999; 99WO-CA00552.

PR 12-JUN-1998; 98US-0096541.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Friedland HD, Krieger TJ, Taylor R, Erfle D, Fraser JR, West MHP;

DR WPI; 2000-223549/19.

Novel pharmaceutical composition containing optionally activated
 polyoxyalkylene-modified cationic peptides, useful for treating tumours
 -
 Disclosure: Page 15; 94pp; English.

This sequence represents a cationic peptide amino acid sequence, which
 can be used in the pharmaceutical composition of the invention. The
 invention relates to a pharmaceutical composition containing at least one
 activated polyoxyalkylene (APO)-modified cationic peptide. The
 modification of peptides with APO increases their activity against tumour
 cells, including those with a multidrug resistant phenotype. The
 pharmaceutical composition can be used to treat tumours, specifically
 lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 cervix, uterus, skin, prostate, liver and colon.

SQ Sequence 20 AA;

Query Match 100.0%; Score 86; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 5,1e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRMPMPWRRK 12
 |||||
 DB 1 ILRMPMPWRRK 12

RESULT 6

AAV24552

AAV24552 standard; peptide: 21 AA.

AAV24552:

18-AUG-1999 (first entry)

Indolicidin analogue #4.

Indolicidin: bacterial infection; photo-oxidised solubilisier; antimicrobial; antibiotic; antiaerhythmic; surface disinfectant; additive; shampoo; soap; insecticide; herbicide; preservative; food; technical material.

Synthetic.

WO9807745-A2.

26-FEB-1998.

21-AUG-1997: 97WO-US14779.

13-JAN-1997: 97US-0034949.

21-AUG-1996: 96US-0024754.

(MICR-) MICROLOGIX BIOTECH INC.

Erle D, Fraser JR, Krieger TJ, Taylor R, West MH; WPI: 1998-169090/15.

New indolicidin analogues with antimicrobial activity and related nucleic acid - vectors, transformed cells and antibodies, also conjugates with polyoxyalkylene glycol and fatty acid to reduce toxicity, useful therapeutically, as disinfectants etc.

Claim 11: Page 88; 129pp: English.

AAV24549 to AAV24615 represent indolicidin analogues of formulae (I)-(VIII) containing up to 25 amino acids (aa): RXZXXXB (I), BXZXXZB (II), BBXZXXZB (III), BXZXXZBBn(AA)nMLBBAGS (IV), BXZXXZBB(AA)nM (V), LBHXXZXXZB (VI), LKXZXXZB (VII) and BBXZXXZBB (VIII). Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa; preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VIII) at least 2 X = F or Y. The analogues are used to treat infections caused by bacteria (Gram positive or negative, or anaerobic); fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or trematodes) or viruses. Typical of very many pathogens that can be controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived from the analogues may be used similarly; the compounds may also be prepared from antibiotics or antiarrhythmic agents; also they are useful as surface disinfectants, as additives to shampoo or soaps, as insecticides or herbicides, or as preservatives for foods and technical materials. The analogues are administered by injection, lavage, orally or topically, generally at 0.1-50 mg/kg. These analogues have a broader spectrum of activity than indolicidin and modification as compounds reduces their toxicity.

Sequence 21 AA:

Query Match 100.0%; Score 86; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPMWPMRRK 12
|||||
DB 1 ILRWPMWPMRRK 12

RESULT 7

AAV24554 standard; peptide: 21 AA.

AAV24554:

18-AUG-1999 (first entry)

Indolicidin analogue #6.

Indolicidin: bacterial infection; photo-oxidised solubilisier; antimicrobial; antibiotic; antiaerhythmic; surface disinfectant; additive; shampoo; soap; insecticide; herbicide; preservative; food; technical material.

Synthetic.

WO9807745-A2.

26-FEB-1998.

21-AUG-1997: 97WO-US14779.

13-JAN-1997: 97US-0034949.

21-AUG-1996: 96US-0024754.

(MICR-) MICROLOGIX BIOTECH INC.

Erle D, Fraser JR, Krieger TJ, Taylor R, West MH; WPI: 1998-169090/15.

New indolicidin analogues with antimicrobial activity and related nucleic acid - vectors, transformed cells and antibodies, also conjugates with polyoxyalkylene glycol and fatty acid to reduce toxicity, useful therapeutically, as disinfectants etc.

Claim 11: Page 88; 129pp: English.

AAV24549 to AAV24615 represent indolicidin analogues of formulae (I)-(VIII) containing up to 25 amino acids (aa): RXZXXXB (I), BXZXXZB (II), BBXZXXZB (III), BXZXXZBBn(AA)nMLBBAGS (IV), BXZXXZBB(AA)nM (V), LBHXXZXXZB (VI), LKXZXXZB (VII) and BBXZXXZBB (VIII). Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa; preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V; in (VIII) at least 2 X = F or Y. The analogues are used to treat infections caused by bacteria (Gram positive or negative, or anaerobic); fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or trematodes) or viruses. Typical of very many pathogens that can be controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds derived from the analogues may be used similarly; the compounds may also be prepared from antibiotics or antiarrhythmic agents; also they are useful as surface disinfectants, as additives to shampoo or soaps, as insecticides or herbicides, or as preservatives for foods and technical materials. The analogues are administered by injection, lavage, orally or topically, generally at 0.1-50 mg/kg. These analogues have a broader spectrum of activity than indolicidin and modification as compounds reduces their toxicity.

Sequence 21 AA:

Query Match 100.0%; Score 86; DB 19; Length 21;
Best Local Similarity 100.0%; Pred. No. 5.4e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPMWPMRRK 12
|||||
DB 1 ILRWPMWPMRRK 12

XX (MICR-) MICROLOGIX BIOTECH INC.
 XX Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 XX Novel pharmaceutical composition containing optionally activated
 XX polyoxalkylene-modified cationic peptides; useful for treating tumours
 XX
 XX Disclosure; Page 15; 94pp; English.
 XX This sequence represents a cationic peptide amino acid sequence, which
 XX can be used in the pharmaceutical composition of the invention. The
 XX invention relates to a pharmaceutical composition containing at least one
 XX activated polyoxalkylene (APO)-modified cationic peptide. The
 XX modification of peptides with APO increases their activity against tumour
 XX cells, including those with a multidrug resistant phenotype. The
 XX pharmaceutical composition can be used to treat tumours, specifically
 XX lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 XX cervix, uterus, skin, prostate, liver and colon.
 XX
 XX Sequence 21 AA;
 XX
 XX Query Match 100.0%; Score 86; DB 21; Length 21;
 XX Best Local Similarity 100.0%; Pred. No. 5.4e-06;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 ILRPMWPMWRK 12
 XX |
 XX 1 ILRPMWPMWRK 12
 XX
 XX DB
 XX
 XX RESULT 11
 XX AAM66363 standard; peptide: 27 AA.
 XX ID AAM66363
 XX AC AAM66363;
 XX XX 12-JAN-1999 (first entry)
 XX DT Indollicidin analogue MBI 11B20.
 XX DE Indollicidin analogue; resistance; cationic peptide; antibiotic;
 XX XX Indollicidin analogue; resistance; cationic peptide; antibiotic;
 XX KM bacterial infection; tolerance; antibacterial; microorganism;
 XX KM bacteria; fungus; parasite; virus.
 XX OS Bcs taurus.
 XX OS Synthetic.
 XX XX WO9840401-A2.
 XX PN 17-SEP-1998.
 XX PD 10-MAR-1998; 98WO-CA00190.
 XX PF 25-FEB-1998; 98US-0030619.
 XX PR 10-MAR-1997; 97US-0040649.
 XX PR 20-AUG-1997; 97US-0915314.
 XX PR 26-SEP-1997; 97US-0060099.
 XX XX (MICR-) MICROLOGIX BIOTECH INC.
 XX XX Fraser JR, McNicol PJ, West MHP;
 XX WPI; 1998-520800/44.
 XX DR New indollicidin peptide analogues - useful for, e.g. enhancing
 XX PT activity of antibiotic or overcoming tolerance, acquired resistance
 XX PT or inherent resistance of microorganisms
 XX XX

PS Claim 1; Page 91; 105pp; English.
 XX The present sequence represents an indollicidin analogue. The present
 XX invention describes compositions and methods for treating infection,
 XX especially bacterial infections. The compositions and methods use
 XX cationic peptides in combination with an antibiotic agent which are
 XX then administered to a patient to enhance the activity of the antibiotic
 XX agent, to overcome: (a) tolerance; (b) acquired resistance; and (c)
 XX inherent resistance. The combinations of antibiotics and cationic
 XX peptides can provide synergistic activity against a microorganism that
 XX is tolerant, inherently resistant, or has acquired resistance to an
 XX antibiotic agent. They can be used for killing e.g. bacteria, fungi,
 XX parasites and viruses.
 XX
 XX Sequence 27 AA;
 XX
 XX Query Match 100.0%; Score 86; DB 19; Length 27;
 XX Best Local Similarity 100.0%; Pred. No. 6.9e-06;
 XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 1 ILRPMWPMWRK 12
 XX |
 XX 1 ILRPMWPMWRK 12
 XX
 XX DB
 XX
 XX RESULT 12
 XX AAY91800 standard; peptide: 28 AA.
 XX ID AAY91800
 XX AC AAY91800;
 XX XX 06-JUN-2000 (first entry)
 XX DT Amino acid sequence of cationic peptide MBI 11B20CN.
 XX DE Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
 XX XX leukaemia; polyoxalkylene-modified; APO; lymphoma; multiple myeloma;
 XX KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
 XX KM multidrug resistance.
 XX OS Synthetic.
 XX OS WO965506-A2.
 XX PN 23-DEC-1999.
 XX PD 14-JUN-1999; 99WO-CA00552.
 XX PF 12-JUN-1998; 98US-0096541.
 XX PR (MICR-) MICROLOGIX BIOTECH INC.
 XX XX Friedland HD, Krieger TJ, Taylor R, Erfile D, Fraser JR, West MHP;
 XX WPI; 2000-223549/19.
 XX DR Novel pharmaceutical composition containing optionally activated
 XX PT polyoxalkylene-modified cationic peptides; useful for treating tumours
 XX PT
 XX
 XX Claim 1; Page 15; 94pp; English.
 XX This sequence represents a cationic peptide amino acid sequence, which
 XX can be used in the pharmaceutical composition of the invention. The
 XX invention relates to a pharmaceutical composition containing at least one
 XX activated polyoxalkylene (APO)-modified cationic peptide. The
 XX modification of peptides with APO increases their activity against tumour
 XX cells, including those with a multidrug resistant phenotype. The
 XX pharmaceutical composition can be used to treat tumours, specifically
 XX lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
 XX cervix, uterus, skin, prostate, liver and colon.
 XX

SQ Sequence 28 AA;

Query Match 100.0%; Score 86; DB 21; Length 28;
 Best Local Similarity 100.0%; Pred. No. 7.1e-06;
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
 |||:|||||
 Db 1 ILRPMWPMRRK 12

RESULT 13

AA24567

ID AAY24567 standard; peptide; 12 AA.

AC AAY24567;

DT 18-AUG-1999 (first entry)

DE Indolicidin analogue #19.

DE Indolicidin: bacterial infection; photo-oxidised solubilisier;

KM antimicrobial; antibiotic; antiarrhythmic; surface disinfectant;

KM additive: shampoo; soap; insecticide; herbicide; preservative;

KM food; technical material.

OS Synthetic.

PN MO9807745-A2.

PD 26-FEB-1998.

PF 21-AUG-1997; 97WO-US14779.

PR 13-JAN-1997; 97US-0034949.

PR 21-AUG-1996; 96US-0024754.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Erfle D, Fraser JR, Krieger TJ, Taylor R, West MH;

DR MPI; 1998-169090/15.

XX New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxyalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.

PS Claim 12; Page 89; 129pp; English.

XX AAY24549 to AAY24615 represent indolicidin analogues of formulae
 CC (I)-(VIII) containing up to 25 amino acids (aa): R₁XXXXX (I), B₁XXXXX
 CC (II), B₁XXXXX (III), B₁XXXXX (IV), B₁XXXXX (V), B₁XXXXX (VI),
 CC (V), B₁XXXXX (VII), L₁XXXXX (VIII) and B₁XXXXX (IX).
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa,
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC fungi (yeast or moulds); parasites (protozoa, nematodes, cestodes or
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and parvovirus. Compounds
 CC derived from the analogues may be used similarly: the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents. The analogues
 CC may be used therapeutically or to coat medical devices; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.

SX Sequence 12 AA;

Query Match 96.5%; Score 83; DB 19; Length 12;
 Best Local Similarity 91.7%; Pred. No. 7.9e-06;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
 |||:|||||
 Db 1 ILRPMWPMRRK 12

RESULT 14

AA91788

ID AAY91788 standard; Peptide; 12 AA.

AC AAY91788;

DT 06-JUN-2000 (first entry)

DE Amino acid sequence of cationic peptide MBI 11B3CN.

KM Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;

KM leukemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;

KM breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;

KM multidrug resistance.

OS Synthetic.

PN MO9965506-A2.

PD 23-DEC-1999.

PF 14-JUN-1999; 99WO-CA00552.

PR 12-JUN-1998; 98US-0096541.

PA (MICR-) MICROLOGIX BIOTECH INC.

PI Friedland HD, Krieger TJ, Taylor R, Fraser JR, West MHP;

DR MPI; 2000-223549/19.

XX Novel pharmaceutical composition containing optionally activated
 PT polyoxyalkylene-modified cationic peptides, useful for treating tumours

PS Disclosure; Page 14; 94pp; English.

XX This sequence represents a cationic peptide amino acid sequence, which
 CC can be used in the pharmaceutical composition of the invention. The
 CC invention relates to a pharmaceutical composition containing at least one
 CC activated polyoxyalkylene (APO)-modified cationic peptide. The
 CC modification of peptides with APO increases their activity against tumour
 CC cells, including those with a multidrug resistant phenotype. The
 CC pharmaceutical composition can be used to treat tumours, specifically
 CC lymphoma, leukemia, multiple myeloma, or tumours of breast, lung, ovary,
 CC cervix, uterus, skin, prostate, liver and colon.

SQ Sequence 12 AA;

Query Match 96.5%; Score 83; DB 21; Length 12;
 Best Local Similarity 91.7%; Pred. No. 7.9e-06;
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
 |||:|||||
 Db 1 ILRPMWPMRRK 12

RESULT 15

Search completed: June 21, 2001, 08:33:12
 Job time: 259 sec

AAV24594
 ID AAV24594 standard; peptide; 12 AA.
 AC AAV24594:
 XX 18-AUG-1999 (first entry)
 DF Indolicidin analogue #46.
 DE Indolicidin analogue #46.
 XX Indolicidin; bacterial infection; photo-oxidised solubiliser;
 XX antimicrobial; antibiotic; antidiarrhetic; surface disinfectant;
 KW additive; shampoo; soap; insecticide; herbicide; preservative;
 KW food; technical material.
 OS Synthetic.
 OS WO9807745-A2.
 PN 26-FEB-1998.
 PD 21-AUG-1997; 97WO-US14779.
 PF 13-JAN-1997; 97US-0034949.
 PR 21-AUG-1996; 96US-0024754.
 PA (MICR-) MICROLOGIX BIOTECH INC.
 XX Erle D, Fraser JR, Krieger TJ, Taylor R, West MH;
 PI WPI; 1998-169090/15.
 DR New indolicidin analogues with antimicrobial activity and related
 PT nucleic acid - vectors, transformed cells and antibodies, also
 PT conjugates with polyoxyalkylene glycol and fatty acid to reduce
 PT toxicity, useful therapeutically, as disinfectants etc.
 PS Claim 14; Page 89; 129pp; English.
 XX AAV24594 to AAV24615 represent indolicidin analogues of formulae
 CC (I)-(VII) containing up to 25 amino acids (aa): RXXXXXB (I), BXXXXXB
 CC (II), BXXXXXB (III), BXXXXBBB(AA)MILBAGS (IV), BXXXXBBB(AA)M
 CC (V), LBHXXNXXNXXR (VI), LKXXNXXNXXR (VII) and BXXXXNXXBB (VIII).
 CC Where Z = P or V; X = hydrophobic residue, preferably W; B = basic aa;
 CC preferably R or K; AA = any aa; n = 0 or 1; in (II), at least 1 Z = V;
 CC in (VIII) at least 2 X = F or Y. The analogues are used to treat
 CC infections caused by bacteria (Gram positive or negative, or anaerobic);
 CC infections caused by parasites (protozoa, nematodes, cestodes or
 CC fungi (yeast or moulds)); typical of very many pathogens that can be
 CC trematodes) or viruses. Typical of very many pathogens that can be
 CC controlled are Leishmania, Trypanosoma, Ascaris lumbricoides, Fasciola
 CC hepatica, Klebsiella pneumoniae, Bordetella pertussis, Staphylococcus
 CC aureus, Listeria, Clostridium, rotavirus and papilloma virus. Compounds
 CC derived from the analogues may be used similarly; the compounds may
 CC also be prepared from antibiotics or antiarrhythmic agents; also they are
 CC useful as surface disinfectants, as additives to shampoo or soaps, as
 CC insecticides or herbicides, or as preservatives for foods and technical
 CC materials. The analogues are administered by injection, lavage, orally
 CC or topically, generally at 0.1-50 mg/kg. These analogues have a broader
 CC spectrum of activity than indolicidin and modification as compounds
 CC reduces their toxicity.
 CC Sequence 12 AA:
 SQ

Query Match 95.3%; Score 82; DB 19; Length 12;
 Best Local Similarity 100.0%; Pred. No. 1,1e-05;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRPWMPWRK 12
 | | | | | | | | | | | | | |
 DB 2 LRPWMPWRK 12

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:33:33 ; Search time 17.69 Seconds
(without alignments)
13.665 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWPMPWRRK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 193259 seqs, 20144635 residues.

Total number of hits satisfying chosen parameters: 193259

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: /cgn2_patents/AA:*
2: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
7: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	12	4	US-08-915-314-42 Sequence 42, Appl
2	86	100.0	20	4	US-08-915-314-47 Sequence 47, Appl
3	86	100.0	21	4	US-08-915-314-46 Sequence 46, Appl
4	86	100.0	21	4	US-08-915-314-48 Sequence 48, Appl
5	83	96.5	12	4	US-08-915-314-40 Sequence 40, Appl
6	82	95.3	12	4	US-08-915-314-76 Sequence 76, Appl
7	81	94.2	12	4	US-08-915-314-77 Sequence 77, Appl
8	81	94.2	12	4	US-08-915-314-87 Sequence 87, Appl
9	80	93.0	12	4	US-08-915-314-78 Sequence 78, Appl
10	80	93.0	12	4	US-08-915-314-85 Sequence 85, Appl
11	80	93.0	12	4	US-08-915-314-86 Sequence 86, Appl
12	78	90.7	12	4	US-08-915-314-80 Sequence 80, Appl
13	78	90.7	12	4	US-08-915-314-83 Sequence 83, Appl
14	78	90.7	13	4	US-08-915-314-38 Sequence 38, Appl
15	77	89.5	12	4	US-08-915-314-69 Sequence 69, Appl
16	77	87.2	11	4	US-08-915-314-41 Sequence 41, Appl
17	75	87.2	12	4	US-08-915-314-52 Sequence 52, Appl
18	75	87.2	13	4	US-08-915-314-25 Sequence 25, Appl
19	75	87.2	13	4	US-08-915-314-30 Sequence 30, Appl
20	75	87.2	13	4	US-08-915-314-51 Sequence 51, Appl
21	75	87.2	13	4	US-08-915-314-62 Sequence 62, Appl
22	75	87.2	13	4	US-08-915-314-63 Sequence 63, Appl
23	75	87.2	13	4	US-08-915-314-64 Sequence 64, Appl
24	75	87.2	13	4	US-08-702-054B-33 Sequence 33, Appl
25	75	87.2	13	4	US-08-702-054B-34 Sequence 34, Appl
26	75	87.2	13	4	US-08-702-054B-35 Sequence 35, Appl
27	75	87.2	14	4	US-08-915-314-57 Sequence 57, Appl

28	75	87.2	16	4	US-08-702-054B-2	Sequence 2, Appl
29	75	87.2	21	4	US-08-915-314-54	Sequence 54, Appl
30	73	84.9	9	4	US-08-915-314-90	Sequence 90, Appl
31	73	84.9	11	4	US-08-915-314-44	Sequence 44, Appl
32	73	84.9	16	4	US-08-702-054B-38	Sequence 38, Appl
33	72	83.7	12	4	US-08-915-314-79	Sequence 79, Appl
34	72	83.7	12	4	US-08-915-314-81	Sequence 81, Appl
35	72	83.7	12	4	US-08-915-314-82	Sequence 82, Appl
36	72	83.7	12	4	US-08-915-314-84	Sequence 84, Appl
37	70	81.4	11	4	US-08-702-054B-9	Sequence 9, Appl
38	70	81.4	12	4	US-08-915-314-39	Sequence 39, Appl
39	70	81.4	12	4	US-08-915-314-74	Sequence 74, Appl
40	70	81.4	12	4	US-08-702-054B-5	Sequence 5, Appl
41	70	81.4	13	1	US-07-715-271-1	Sequence 1, Appl
42	70	81.4	13	1	US-08-197-205-1	Sequence 1, Appl
43	70	81.4	13	1	US-08-197-205-5	Sequence 5, Appl
44	70	81.4	13	4	US-08-915-314-29	Sequence 29, Appl
45	70	81.4	13	4	US-08-915-314-58	Sequence 58, Appl

ALIGNMENTS

RESULT 1
US-08-915-314-42
Sequence 42, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: NO. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-42

Query Match 100.0%; Score 86; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 1.2e-06;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ILRWPMPWRRK 12

DB 1 ILRMPWMPWRRK 12

RESULT 2

US-08-915-314-47
Sequence 47, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/915,314
APPLICATION NUMBER: 660081.405
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-47

Query Match 100.0%; Score 86; DB 4; Length 20;

Best Local Similarity 100.0%; Pred. No. 2e-06; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRMPWMPWRRK 12

DB 1 ILRMPWMPWRRK 12

RESULT 3

US-08-915-314-46
Sequence 46, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/915,314
APPLICATION NUMBER: 660081.405
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: No. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 682-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-46

Query Match 100.0%; Score 86; DB 4; Length 21;

Best Local Similarity 100.0%; Pred. No. 2.1e-06; Indels 0; Gaps 0;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRMPWMPWRRK 12

DB 1 ILRMPWMPWRRK 12

RESULT 4

US-08-915-314-48
Sequence 48, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/915,314
APPLICATION NUMBER: 660081.405
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-48

Query Match 100.0%; Score 86; DB 4; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.1e-06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRPMWPMWRRK 12
|||
Db 1 ILRPMWPMWRRK 12

RESULT 5
US-08-915-314-40
Sequence 40, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-40

Query Match 96.5%; Score 83; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 3.1e-06;
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ILRPMWPMWRRK 12
|||
Db 1 ILRPMWPMWRRK 12

RESULT 6
US-08-915-314-76
Sequence 76, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-76

Query Match 95.3%; Score 82; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 4.3e-06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LRMWPMWRRK 12
|||
Db 2 LRMWPMWRRK 12

RESULT 7
US-08-915-314-77
Sequence 77, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erfle, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington

COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 77:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-77

Query Match 94.2% Score 81; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 5.9e-06;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
| | | | | | | | | | | | | | | |
DB 1 IARPMWPMRRK 12

RESULT 8
US-08-915-314-87
Sequence 87, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
NUMBER OF SEQUENCES: 90
INFECTIONS USING ANALOGUES OF INDOLICIDIN
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 87:

SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-87

Query Match 94.2% Score 81; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.9e-06;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ILRPMWPMRR 11
| | | | | | | | | | | | | | | |
DB 1 ILRPMWPMRR 11

RESULT 9
US-08-915-314-78
Sequence 78, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
NUMBER OF SEQUENCES: 90
INFECTIONS USING ANALOGUES OF INDOLICIDIN
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 78:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-78

Query Match 93.0% Score 80; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 8.2e-06;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ILRPMWPMRRK 12
| | | | | | | | | | | | | | | |
DB 1 ILRPMWPMRRK 12

RESULT 10
US-08-915-314-85

Sequence 85, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 85:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-85

Query Match 93.0%; Score 80; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 8.2e-06;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILRPMWPMWRK 12
|1111111111|
Db 1 ILRPMWPMWRK 12

RESULT 11
US-08-915-314-86
Sequence 86, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-86

Query Match 93.0%; Score 80; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 8.2e-06;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILRPMWPMWRK 12
|1111111111|
Db 1 ILRPMWPMWRK 12

RESULT 12
US-08-915-314-80
Sequence 80, Application US/08915314
Patent No. 6180604
GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 80:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
US-08-915-314-80

Query Match 90.7%; Score 78; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 1.6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILRPMWPMRKK 12
DB 1 ILRPMWPMRKK 12

RESULT 13
US-08-915-314-83
Sequence 83, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:
APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 83:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-83

Query Match 90.7%; Score 78; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 1.6e-05;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ILRPMWPMRKK 12
DB 1 ILRPMWPMRKK 12

RESULT 14
US-08-915-314-38
Sequence 38, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/915,314
FILING DATE: 20-AUG-1997

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: No. 6180604tendburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 660081.405
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 38:

SEQUENCE CHARACTERISTICS:

LENGTH: 13 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-915-314-38

Query Match 90.7%; Score 78; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RMPWPMRKK 12
DB 4 RMPWPMRKK 13

RESULT 15
US-08-915-314-69
Sequence 69, Application US/08915314
Patent No. 6180604

GENERAL INFORMATION:

APPLICANT: Fraser, Janet R.
APPLICANT: West, Michael H.P.
APPLICANT: Krieger, Timothy J.
APPLICANT: Taylor, Robert
APPLICANT: Erile, Douglas

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
TITLE OF INVENTION: INFECTIONS USING ANALOGUES OF INDOLICIDIN
NUMBER OF SEQUENCES: 90
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/915,314
: FILING DATE: 20-AUG-1997
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: No. 6180604tenburg Ph.D., Carol
: REGISTRATION NUMBER: 39,317
: REFERENCE/DOCKET NUMBER: 660081.405
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 69:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 12 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: Linear
: US-08-915-314-69

```

```

Query Match      89.5%; Score 77; DB 4; Length 12;
Best Local Similarity 83.3%; Pred. No. 2,1e-05;
Matches 10; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      1 ILRMPMPWRRK 12
      . | : |||||
Db      1 IKKMPMPWRRK 12

```

Search completed: June 21, 2001, 08:33:33
 Job time: 175 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:33:57 ; Search time 20.32 Seconds
(without alignments)
44.985 Million cell updates/sec

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRMPMPWRK 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1.*
2: p1r2.*
3: p1r3.*
4: p1r4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	81.4	144	1	JC1222
2	53	61.6	1173	1	VGIRHC
3	51	59.3	299	2	T12505
4	50	58.1	111	2	T29295
5	49	57.0	498	1	JT0751
6	48.5	56.4	114	2	T36208
7	47	54.7	248	2	S23449
8	47	54.7	253	2	G70715
9	47	54.7	276	2	B83161
10	47	54.7	715	2	B70741
11	47	54.7	1411	2	T48529
12	46	53.5	728	2	T51071
13	45.5	52.9	505	2	A39128
14	45	52.3	196	2	S55483
15	45	52.3	273	2	F82646
16	45	52.3	412	2	A83604
17	45	52.3	448	2	H72376
18	45	52.3	1108	2	A48508
19	44	51.2	257	2	S70177
20	44	51.2	361	2	A36669
21	44	51.2	397	2	B70763
22	44	51.2	535	2	T38244
23	44	51.2	621	2	S37664
24	44	51.2	630	2	S37663
25	44	51.2	967	2	C70831
26	44	51.2	968	2	F70746
27	44	51.2	968	2	T00322
28	44	51.2	1154	1	VGIRHB
29	44	51.2	1162	1	VGIRAK

30	44	51.2	1162	2	S07421	E2 glycoprotein pr
31	44	51.2	1162	2	S14939	E2 glycoprotein pr
32	44	51.2	1162	2	S14940	E2 glycoprotein pr
33	43.5	50.6	1529	2	A59189	Arp-binding casset
34	43	50.0	51	2	S23291	light-harvesting p
35	43	50.0	192	2	H86543	hypothetical prote
36	43	50.0	192	2	D72081	conserved hypothet
37	43	50.0	236	2	J00606	arylesterase (EC 3
38	43	50.0	250	2	A83506	probable cobalamin
39	43	50.0	278	2	T46458	hypothetical prote
40	43	50.0	298	2	B72492	hypothetical prote
41	43	50.0	646	2	H82555	c-type cytochrome
42	43	50.0	711	2	C40046	antibiotic transpo
43	43	50.0	738	2	F96701	hypothetical prote
44	43	50.0	958	2	A70634	probable mmp1 pro
45	43	50.0	1112	2	S70522	cyclic nucleotide

ALIGNMENTS

RESULT 1

JC1222

Indolicidin precursor - bovine

N:Alternate names: antimicrobial peptide

C:Species: Bos primigenius taurus (cattle)

C>Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999

C:Accession: JC1222; A42387; S25664

R:del Sal, G.; Storici, P.; Schneider, C.; Romeo, D.; Zanetti, M.

Biochem. Biophys. Res. Commun. 187, 467-472, 1992

A:Title: CDNA cloning of the neutrophil bactericidal peptide indolicidin.

A:Reference number: JC1222; MUID:92392368

A:Accession: JC1222

A:Molecule type: mRNA

A:Residues: 1-144 <SRL>

A:Cross-References: EMBL:X67340; NID:9462; PIDN:CAA47755.1; PID:9463

A:Experimental source: bone marrow

R:Seasted, M.E.; Novotny, M.J.; Morris, W.L.; Tang, Y.Q.; Smith, W.; Cullor, J.S.

J. Biol. Chem. 267, 4292-4295, 1992

A:Title: Indolicidin, a novel bactericidal tridecapeptide amide from neutrophils.

A:Reference number: A42387; MUID:92165771

A:Accession: A42387

A:Molecule type: protein

A:Residues: 131-143 <SRL>

A:Experimental source: neutrophils

A:Note: sequence extracted from NCBI backbone (NCBIP:83840)

C:Superfamily: cathelin; cystatin homology

C:Keywords: amidated carboxyl end

F:1-29/Domain: signal sequence #status predicted <SIG>

F:122-129/Domain: cystatin homology <CYS>

F:30-130/Domain: propeptide #status predicted <PRO>

F:31-143/Product: indolicidin #status experimental <MAT>

F:143/Modified site: amidated carboxyl end (Arg) (amide in mature form from followi

Query Match

Best Local Similarity

Matches 8; Conservative

OY 3 RMPMPWR 11

DB 135 KMPMPWR 143

RESULT 2

VGIRHC

E2 glycoprotein precursor - human coronavirus (strain 229E)

N:Alternate names: peplomer glycoprotein; s glycoprotein; spike glycoprotein

C:Species: human coronavirus

A:Note: host Homo sapiens (man)

C>Date: 31-Dec-1991 #sequence-revision 31-Dec-1991 #text-change 16-Jun-2000

C:Accession: A34766; S05460

R:Rabe, T.; Schelle-Prinz, B.; Sidell, S.G.

J. Gen. Virol. 71, 1065-1073, 1990
 A:Title: Nucleotide sequence of the gene encoding the spike glycoprotein of human corona
 A:Reference number: A34766; MUID:90264837
 A:Accession: A34766
 A:Molecule type: mRNA
 A:Residues: 1-1173 <RAA>
 A:Cross-references: EMBL:X16816; NID:958926; PIDN:CAA34723.1; PID:958927
 A:Experimental source: strain 229E
 R:Rabe, T.; Stidell, S.
 Nucleic Acids Res. 17, 6387, 1989
 A:Title: Nucleotide sequence of the human coronavirus HCV 229E mRNA 4 and mRNA 5 unique
 A:Reference number: A34038; MUID:89366667
 A:Accession: S05460
 A:Status: translation not shown
 A:Molecule type: mRNA
 A:Residues: 1159-1173 <RA2>
 A:Cross-references: EMBL:X15654; NID:958921; PIDN:CAA33680.1; PID:91334827
 C:Superfamily: coronavirus E2 glycoprotein
 C:Keywords: glycoprotein; transmembrane protein
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-1173/Product: E2 glycoprotein #status predicted <MAN>
 F:116-1136/Domain: transmembrane #status predicted <MAN>
 F:23,62,98,147,171,176,220,243,326,333,440,464,518,538,568,581,587,663,671,930,1015,
 1119,1120,1121,1122,1123,1124,1125,1126,1127,1128,1129,1130,1131,1132,1133,1134,1135,1136,1137,1138,1139,1140,1141,1142,1143,1144,1145,1146,1147,1148,1149,1150,1151,1152,1153,1154,1155,1156,1157,1158,1159,1160,1161,1162,1163,1164,1165,1166,1167,1168,1169,1170,1171,1172,1173,1174,1175,1176,1177,1178,1179,1180,1181,1182,1183,1184,1185,1186,1187,1188,1189,1190,1191,1192,1193,1194,1195,1196,1197,1198,1199,1200,1201,1202,1203,1204,1205,1206,1207,1208,1209,1210,1211,1212,1213,1214,1215,1216,1217,1218,1219,1220,1221,1222,1223,1224,1225,1226,1227,1228,1229,1230,1231,1232,1233,1234,1235,1236,1237,1238,1239,1240,1241,1242,1243,1244,1245,1246,1247,1248,1249,1250,1251,1252,1253,1254,1255,1256,1257,1258,1259,1260,1261,1262,1263,1264,1265,1266,1267,1268,1269,1270,1271,1272,1273,1274,1275,1276,1277,1278,1279,1280,1281,1282,1283,1284,1285,1286,1287,1288,1289,1290,1291,1292,1293,1294,1295,1296,1297,1298,1299,1300,1301,1302,1303,1304,1305,1306,1307,1308,1309,1310,1311,1312,1313,1314,1315,1316,1317,1318,1319,1320,1321,1322,1323,1324,1325,1326,1327,1328,1329,1330,1331,1332,1333,1334,1335,1336,1337,1338,1339,1340,1341,1342,1343,1344,1345,1346,1347,1348,1349,1350,1351,1352,1353,1354,1355,1356,1357,1358,1359,1360,1361,1362,1363,1364,1365,1366,1367,1368,1369,1370,1371,1372,1373,1374,1375,1376,1377,1378,1379,1380,1381,1382,1383,1384,1385,1386,1387,1388,1389,1390,1391,1392,1393,1394,1395,1396,1397,1398,1399,1400,1401,1402,1403,1404,1405,1406,1407,1408,1409,1410,1411,1412,1413,1414,1415,1416,1417,1418,1419,1420,1421,1422,1423,1424,1425,1426,1427,1428,1429,1430,1431,1432,1433,1434,1435,1436,1437,1438,1439,1440,1441,1442,1443,1444,1445,1446,1447,1448,1449,1450,1451,1452,1453,1454,1455,1456,1457,1458,1459,1460,1461,1462,1463,1464,1465,1466,1467,1468,1469,1470,1471,1472,1473,1474,1475,1476,1477,1478,1479,1480,1481,1482,1483,1484,1485,1486,1487,1488,1489,1490,1491,1492,1493,1494,1495,1496,1497,1498,1499,1500,1501,1502,1503,1504,1505,1506,1507,1508,1509,1510,1511,1512,1513,1514,1515,1516,1517,1518,1519,1520,1521,1522,1523,1524,1525,1526,1527,1528,1529,1530,1531,1532,1533,1534,1535,1536,1537,1538,1539,1540,1541,1542,1543,1544,1545,1546,1547,1548,1549,1550,1551,1552,1553,1554,1555,1556,1557,1558,1559,1560,1561,1562,1563,1564,1565,1566,1567,1568,1569,1570,1571,1572,1573,1574,1575,1576,1577,1578,1579,1580,1581,1582,1583,1584,1585,1586,1587,1588,1589,1590,1591,1592,1593,1594,1595,1596,1597,1598,1599,1600,1601,1602,1603,1604,1605,1606,1607,1608,1609,1610,1611,1612,1613,1614,1615,1616,1617,1618,1619,1620,1621,1622,1623,1624,1625,1626,1627,1628,1629,1630,1631,1632,1633,1634,1635,1636,1637,1638,1639,1640,1641,1642,1643,1644,1645,1646,1647,1648,1649,1650,1651,1652,1653,1654,1655,1656,1657,1658,1659,1660,1661,1662,1663,1664,1665,1666,1667,1668,1669,1670,1671,1672,1673,1674,1675,1676,1677,1678,1679,1680,1681,1682,1683,1684,1685,1686,1687,1688,1689,1690,1691,1692,1693,1694,1695,1696,1697,1698,1699,1700,1701,1702,1703,1704,1705,1706,1707,1708,1709,1710,1711,1712,1713,1714,1715,1716,1717,1718,1719,1720,1721,1722,1723,1724,1725,1726,1727,1728,1729,1730,1731,1732,1733,1734,1735,1736,1737,1738,1739,1740,1741,1742,1743,1744,1745,1746,1747,1748,1749,1750,1751,1752,1753,1754,1755,1756,1757,1758,1759,1760,1761,1762,1763,1764,1765,1766,1767,1768,1769,1770,1771,1772,1773,1774,1775,1776,1777,1778,1779,1780,1781,1782,1783,1784,1785,1786,1787,1788,1789,1790,1791,1792,1793,1794,1795,1796,1797,1798,1799,1800,1801,1802,1803,1804,1805,1806,1807,1808,1809,1810,1811,1812,1813,1814,1815,1816,1817,1818,1819,1820,1821,1822,1823,1824,1825,1826,1827,1828,1829,1830,1831,1832,1833,1834,1835,1836,1837,1838,1839,1840,1841,1842,1843,1844,1845,1846,1847,1848,1849,1850,1851,1852,1853,1854,1855,1856,1857,1858,1859,1860,1861,1862,1863,1864,1865,1866,1867,1868,1869,1870,1871,1872,1873,1874,1875,1876,1877,1878,1879,1880,1881,1882,1883,1884,1885,1886,1887,1888,1889,1890,1891,1892,1893,1894,1895,1896,1897,1898,1899,1900,1901,1902,1903,1904,1905,1906,1907,1908,1909,1910,1911,1912,1913,1914,1915,1916,1917,1918,1919,1920,1921,1922,1923,1924,1925,1926,1927,1928,1929,1930,1931,1932,1933,1934,1935,1936,1937,1938,1939,1940,1941,1942,1943,1944,1945,1946,1947,1948,1949,1950,1951,1952,1953,1954,1955,1956,1957,1958,1959,1960,1961,1962,1963,1964,1965,1966,1967,1968,1969,1970,1971,1972,1973,1974,1975,1976,1977,1978,1979,1980,1981,1982,1983,1984,1985,1986,1987,1988,1989,1990,1991,1992,1993,1994,1995,1996,1997,1998,1999,2000,2001,2002,2003,2004,2005,2006,2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021,2022,2023,2024,2025,2026,2027,2028,2029,2030,2031,2032,2033,2034,2035,2036,2037,2038,2039,2040,2041,2042,2043,2044,2045,2046,2047,2048,2049,2050,2051,2052,2053,2054,2055,2056,2057,2058,2059,2060,2061,2062,2063,2064,2065,2066,2067,2068,2069,2070,2071,2072,2073,2074,2075,2076,2077,2078,2079,2080,2081,2082,2083,2084,2085,2086,2087,2088,2089,2090,2091,2092,2093,2094,2095,2096,2097,2098,2099,2100,2101,2102,2103,2104,2105,2106,2107,2108,2109,2110,2111,2112,2113,2114,2115,2116,2117,2118,2119,2120,2121,2122,2123,2124,2125,2126,2127,2128,2129,2130,2131,2132,2133,2134,2135,2136,2137,2138,2139,2140,2141,2142,2143,2144,2145,2146,2147,2148,2149,2150,2151,2152,2153,2154,2155,2156,2157,2158,2159,2160,2161,2162,2163,2164,2165,2166,2167,2168,2169,2170,2171,2172,2173,2174,2175,2176,2177,2178,2179,2180,2181,2182,2183,2184,2185,2186,2187,2188,2189,2190,2191,2192,2193,2194,2195,2196,2197,2198,2199,2200,2201,2202,2203,2204,2205,2206,2207,2208,2209,2210,2211,2212,2213,2214,2215,2216,2217,2218,2219,2220,2221,2222,2223,2224,2225,2226,2227,2228,2229,2230,2231,2232,2233,2234,2235,2236,2237,2238,2239,2240,2241,2242,2243,2244,2245,2246,2247,2248,2249,2250,2251,2252,2253,2254,2255,2256,2257,2258,2259,2260,2261,2262,2263,2264,2265,2266,2267,2268,2269,2270,2271,2272,2273,2274,2275,2276,2277,2278,2279,2280,2281,2282,2283,2284,2285,2286,2287,2288,2289,2290,2291,2292,2293,2294,2295,2296,2297,2298,2299,2300,2301,2302,2303,2304,2305,2306,2307,2308,2309,2310,2311,2312,2313,2314,2315,2316,2317,2318,2319,2320,2321,2322,2323,2324,2325,2326,2327,2328,2329,2330,2331,2332,2333,2334,2335,2336,2337,2338,2339,2340,2341,2342,2343,2344,2345,2346,2347,2348,2349,2350,2351,2352,2353,2354,2355,2356,2357,2358,2359,2360,2361,2362,2363,2364,2365,2366,2367,2368,2369,2370,2371,2372,2373,2374,2375,2376,2377,2378,2379,2380,2381,2382,2383,2384,2385,2386,2387,2388,2389,2390,2391,2392,2393,2394,2395,2396,2397,2398,2399,2400,2401,2402,2403,2404,2405,2406,2407,2408,2409,2410,2411,2412,2413,2414,2415,2416,2417,2418,2419,2420,2421,2422,2423,2424,2425,2426,2427,2428,2429,2430,2431,2432,2433,2434,2435,2436,2437,2438,2439,2440,2441,2442,2443,2444,2445,2446,2447,2448,2449,2450,2451,2452,2453,2454,2455,2456,2457,2458,2459,2460,2461,2462,2463,2464,2465,2466,2467,2468,2469,2470,2471,2472,2473,2474,2475,2476,2477,2478,2479,2480,2481,2482,2483,2484,2485,2486,2487,2488,2489,2490,2491,2492,2493,2494,2495,2496,2497,2498,2499,2500,2501,2502,2503,2504,2505,2506,2507,2508,2509,2510,2511,2512,2513,2514,2515,2516,2517,2518,2519,2520,2521,2522,2523,2524,2525,2526,2527,2528,2529,2530,2531,2532,2533,2534,2535,2536,2537,2538,2539,2540,2541,2542,2543,2544,2545,2546,2547,2548,2549,2550,2551,2552,2553,2554,2555,2556,2557,2558,2559,2560,2561,2562,2563,2564,2565,2566,2567,2568,2569,2570,2571,2572,2573,2574,2575,2576,2577,2578,2579,2580,2581,2582,2583,2584,2585,2586,2587,2588,2589,2590,2591,2592,2593,2594,2595,2596,2597,2598,2599,2600,2601,2602,2603,2604,2605,2606,2607,2608,2609,2610,2611,2612,2613,2614,2615,2616,2617,2618,2619,2620,2621,2622,2623,2624,2625,2626,2627,2628,2629,2630,2631,2632,2633,2634,2635,2636,2637,2638,2639,2640,2641,2642,2643,2644,2645,2646,2647,2648,2649,2650,2651,2652,2653,2654,2655,2656,2657,2658,2659,2660,2661,2662,2663,2664,2665,2666,2667,2668,2669,2670,2671,2672,2673,2674,2675,2676,2677,2678,2679,2680,2681,2682,2683,2684,2685,2686,2687,2688,2689,2690,2691,2692,2693,2694,2695,2696,2697,2698,2699,2700,2701,2702,2703,2704,2705,2706,2707,2708,2709,2710,2711,2712,2713,2714,2715,2716,2717,2718,2719,2720,2721,2722,2723,2724,2725,2726,2727,2728,2729,2730,2731,2732,2733,2734,2735,2736,2737,2738,2739,2740,2741,2742,2743,2744,2745,2746,2747,2748,2749,2750,2751,2752,2753,2754,2755,2756,2757,2758,2759,2760,2761,2762,2763,2764,2765,2766,2767,2768,2769,2770,2771,2772,2773,2774,2775,2776,2777,2778,2779,2780,2781,2782,2783,2784,2785,2786,2787,2788,2789,2790,2791,2792,2793,2794,2795,2796,2797,2798,2799,2800,2801,2802,2803,2804,2805,2806,2807,2808,2809,2810,2811,2812,2813,2814,2815,2816,2817,2818,2819,2820,2821,2822,2823,2824,2825,2826,2827,2828,2829,2830,2831,2832,2833,2834,2835,2836,2837,2838,2839,2840,2841,2842,2843,2844,2845,2846,2847,2848,2849,2850,2851,2852,2853,2854,2855,2856,2857,2858,2859,2860,2861,2862,2863,2864,2865,2866,2867,2868,2869,2870,2871,2872,2873,2874,2875,2876,2877,2878,2879,2880,2881,2882,2883,2884,2885,2886,2887,2888,2889,2890,2891,2892,2893,2894,2895,2896,2897,2898,2899,2900,2901,2902,2903,2904,2905,2906,2907,2908,2909,2910,2911,2912,2913,2914,2915,2916,2917,2918,2919,2920,2921,2922,2923,2924,2925,2926,2927,2928,2929,2930,2931,2932,2933,2934,2935,2936,2937,2938,2939,2940,2941,2942,2943,2944,2945,2946,2947,2948,2949,2950,2951,2952,2953,2954,2955,2956,2957,2958,2959,2960,2961,2962,2963,2964,2965,2966,2967,2968,2969,2970,2971,2972,2973,2974,2975,2976,2977,2978,2979,2980,2981,2982,2983,2984,2985,2986,2987,2988,2989,2990,2991,2992,2993,2994,2995,2996,2997,2998,2999,3000,3001,3002,3003,3004,3005,3006,3007,3008,3009,3010,3011,3012,3013,3014,3015,3016,3017,3018,3019,3020,3021,3022,3023,3024,3025,3026,3027,3028,3029,3030,3031,3032,3033,3034,3035,3036,3037,3038,3039,3040,3041,3042,3043,3044,3045,3046,3047,3048,3049,3050,3051,3052,3053,3054,3055,3056,3057,3058,3059,3060,3061,3062,3063,3064,3065,3066,3067,3068,3069,3070,3071,3072,3073,3074,3075,3076,3077,3078,3079,3080,3081,3082,3083,3084,3085,3086,3087,3088,3089,3090,3091,3092,3093,3094,3095,3096,3097,3098,3099,3100,3101,3102,3103,3104,3105,3106,3107,3108,3109,3110,3111,3112,3113,3114,3115,3116,3117,3118,3119,3120,3121,3122,3123,3124,3125,3126,3127,3128,3129,3130,3131,3132,3133,3134,3135,3136,3137,3138,3139,3140,3141,3142,3143,3144,3145,3146,3147,3148,3149,3150,3151,3152,3153,3154,3155,3156,3157,3158,3159,3160,3161,3162,3163,3164,3165,3166,3167,3168,3169,3170,3171,3172,3173,3174,3175,3176,3177,3178,3179,3180,3181,3182,3183,3184,3185,3186,3187,3188,3189,3190,3191,3192,3193,3194,3195,3196,3197,3198,3199,3200,3201,3202,3203,3204,3205,3206,3207,3208,3209,3210,3211,3212,3213,3214,3215,3216,3217,3218,3219,3220,3221,3222,3223,3224,3225,3226,3227,3228,3229,3230,3231,3232,3233,3234,3235,3236,3237,3238,3239,3240,3241,3242,3243,3244,3245,3246,3247,3248,3249,3250,3251,3252,3253,3254,3255,3256,3257,3258,3259,3260,3261,3262,3263,3264,3265,3266,3267,3268,3269,3270,3271,3272,3273,3274,3275,3276,3277,3278,3279,3280,3281,3282,3283,3284,3285,3286,3287,3288,3289,3290,3291,3292,3293,3294,3295,3296,3297,3298,3299,3300,3301,3302,3303,3304,3305,3306,3307,3308,3309,3310,3311,3312,3313,3314,3315,3316,3317,3318,3319,3320,3321,3322,3323,3324,3325,3326,3327,3328,3329,3330,3331,3332,3333,3334,3335,3336,3337,3338,3339,3340,3341,3342,3343,3344,3345,3346,3347,3348,3349,3350,3351,3352,3353,3354,3355,3356,3357,3358,3359,3360,3361,3362,3363,3364,3365,3366,3367,3368,3369,3370,3371,3372,3373,3374,3375,3376,3377,3378,3379,3380,3381,3382,3383,3384,3385,3386,3387,3388,3389,3390,3391,3392,3393,3394,3395,3396,3397,3398,3399,3400,3401,3402,3403,3404,3405,3406,3407,3408,3409,3410,3411,3412,3413,3414,3415,3416,3417,3418,3419,3420,3421,3422,3423,3424,3425,3426,3427,3428,3429,3430,3431,3432,3433,3434,3435,3436,3437,3438,3439,3440,

A:Residues: 1-76,'R',78-80,'VWLAITTPSRML',95-123,'RVRT',129-204,211-273,'R',275-322,
A:Cross-references: GB:M17029, NID:q162628; PIDN:AAA30362.1; PID:q162629
A:Experimental source: adrenal cortex
R:Marburton, R.J.; Seybert, D.W.
Biochim. Biophys. Acta 1246, 39-46, 1995
A:Title: Structural and functional characterization of bovine adrenodoxin reductase by 1
A:Reference number: S52100; MUID:95110846
A:Accession: S52100
A:Status: preliminary
A:Molecule type: protein
A:Residues: 'X',34-41,'X',43-48,'X',50-51,304-306,'X',308-309,'X',311-326 <NAR>
C:Comment: Ferredoxin--NADP+ reductase is localized in the matrix of adrenal cortex mito
ferredoxin--NADP+ reductase, adrenodoxin and two forms of cytochrome P-450.
C:Genetics:
A:introns: 27/1; 59/3; 91/3; 132/3; 170/3; 204/3; 246/3; 275/1; 341/3; 399/1; 456/1
C:Function:
A:Description: catalyzes the reversible reduction of NADP+ by reduced ferredoxin or redu
C:Superfamily: human ferredoxin--NADP+ reductase
C:Keywords: alternative splicing; flavoprotein; mitochondrion; monomer; NADP; oxidoreduc
F:1-32/Domain: transit peptide (mitochondrion) #status predicted <SIG>
F:33-498/Product: ferredoxin--NADP+ reductase, long form #status predicted <MAT>
F:33-204,211-498/Product: ferredoxin--NADP+ reductase, short form #status experimental <
F:40-70/Region: beta-alpha-beta FAD nucleotide-binding fold
F:180-190/Region: NADP binding #status predicted
F:281/Binding site: substrate (Lys) #status experimental

Query Match 57.0%; Score 49; DB 1; Length 498;
Best Local Similarity 83.3%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 WPMWPMW 9
Db 6 WPMWPMW 11

RESULT 6
T36208
hypothetical protein SCE36.09 - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T36208
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z21601
A:Accession: T36208
A:Status: preliminary; translated from GB/EMBL/DDAB
A:Molecule type: DNA
A:Residues: 1-114 <OLI>
A:Cross-references: EMBL:AL049763; PIDN:CAB42078.1; GSPDB:GN00070; SCOEDB:SCE36.09
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SCE36.09

Query Match 56.4%; Score 48.5; DB 2; Length 114;
Best Local Similarity 80.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 3 RW-PWMPWR 11
Db 103 RWMPWMPWR 112

RESULT 7
S23449
NADH oxidase (H2O2-forming) (EC 1.6.-.-) - Thermus aquaticus
C:Species: Thermus aquaticus
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 23-Mar-1993
C:Accession: S23449; S24556
R:Park, H.J.; Kreutzer, R.; Reiser, C.O.A.; Sprinzl, M.
Eur. J. Biochem. 205, 875-879, 1992
A:Title: Molecular cloning and nucleotide sequence of the gene encoding a H(2)O(2)-formi

A:Reference number: S23449; MUID:92249331
A:Accession: S23449
A:Molecule type: DNA
A:Residues: 1-248 <PAR>
A:Cross-references: EMBL:X60110
A:Accession: S24556
A:Molecule type: protein
A:Residues: 1-32 <PAR1>
C:Genetics:
A:Gene: nox
C:Keywords: NAD; oxidoreductase
F:1-248/Product: NADH oxidase (H2O2-forming) #status experimental <MAT>

Query Match 54.7%; Score 47; DB 2; Length 248;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PMPWPMW 9
Db 179 PMPWPMW 183

RESULT 8
G70715
hypothetical protein RV0945 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70715
R:Collé, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gord
Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroy
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete ge
A:Reference number: A70500; MUID:98295987
A:Accession: G70715
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-253 <COL>
A:Cross-references: GB:779700; GB:AL123456; NID:g3261628; PIDN:CAB02005.1; PID:q152
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0945
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F:8-190/Domain: short-chain alcohol dehydrogenase homology <SABP>

Query Match 54.7%; Score 47; DB 2; Length 253;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PMPWPMW 9
Db 230 PMPWPMW 234

RESULT 9
B83161
probable short-chain dehydrogenase PA3883 [Imported] - Pseudomonas aeruginosa (stra
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83161
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardy, K.;
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic
A:Reference number: A82950; MUID:20437337
A:Accession: B83161
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-276 <SRO>

A:Cross-references: GB:AE004805; GB:AE004091; NID:9950055; PIDN:AA07270.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA383
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 54.7%; Score 47; DB 2; Length 276;
 Best Local Similarity 70.0%; Pred. No. 18;
 Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 3 RMPMPMPRRK 12
 1 11111111
 DB 197 RSPMPPLRRQ 206

RESULT 10
 B70741
 Probable moey protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis
 C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: B70741
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987
 A:Accession: B70741
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-715 <COI>
 A:Cross-references: GB:Z75555; GB:AL123456; NID:93261608; PIDN:CAA99988.1; PID:e250356;
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: moey

Query Match 54.7%; Score 47; DB 2; Length 715;
 Best Local Similarity 66.7%; Pred. No. 46;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RMPMPMPRR 11
 1 11111111
 DB 65 RWAYTPMR 73

RESULT 11
 T48529
 Hypothetical protein T22P22.90 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T48529
 R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z24490
 A:Accession: T48529
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1411 <BEV>
 A:Cross-references: EMBL:AL163814
 A:Experimental source: cultivar Columbia; BAC clone T22P22
 C:Genetics:
 A:Map position: 5
 A:Insertions: 281/2; 320/1; 389/3; 429/3; 473/3; 515/3; 534/2; 567/3; 602/1; 669/1; 776/2;
 A>Note: T22P22.90

Query Match 54.7%; Score 47; DB 2; Length 1411;
 Best Local Similarity 63.6%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 LRMPMPMPRRK 12
 1 11111111
 DB 1013 LAMSMQOMRRK 1023

RESULT 12
 T51071
 Related to trfA protein [Imported] - Neurospora crassa

N:Alternate names: protein B2A19.50
 C:Species: Neurospora crassa
 C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 21-Jul-2000
 C:Accession: T51071
 R:Schulte, U.; Algn, V.; Honkeisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatu submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25286
 A:Accession: T51071
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <SCH>
 A:Cross-references: EMBL:AL390092; GSPDB:GN00116; NCSP:B2A19.50
 A:Experimental source: BAC clone B2A19; strain OR74A
 C:Genetics:
 A:Gene: NCSP:B2A19.50
 A:Map position: 6
 A:Insertions: 26/1; 119/2

Query Match 53.5%; Score 46; DB 2; Length 728;
 Best Local Similarity 58.3%; Pred. No. 62;
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 ILRMPMPMPRRK 12
 1 11111111
 DB 11 ILGMPMLMWSRR 22

RESULT 13
 A39128
 Anthranilate synthase (EC 4.1.3.27) component I [validated] - Pseudomonas syringae pv

N:Alternate names: anthranilate synthase alpha chain
 C:Species: Pseudomonas syringae pv. savastanoi
 C>Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 17-Mar-2000
 C:Accession: A39128
 R:da Costa, E.; Silva, O.; Kosuge, T.
 J. Bacteriol. 173, 463-471, 1991
 A:Title: Molecular characterization and expression analysis of the anthranilate synth

A:Reference number: A39128; MUID:91100331
 A:Accession: A39128
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-505 <DAC>
 A:Cross-references: GB:M55911
 C:Genetics:
 A:Gene: trpE

C:Complex: heterotetramer; two component I chains, two component II chains
 C:Function: <ANT>
 A:Description: EC 4.1.3.27 [validated; MUID:90130325]
 A:Pathway: tryptophan biosynthesis

A>Note: first step
 C:Function: <COMI>
 A:Description: EC 4.1.3.27 [validated; MUID:91100331]
 A>Note: expression of trpE seems to be independent of the concentration of tryptophan
 C:Superfamily: anthranilate synthase component I
 C:Keywords: carbon-carbon lyase; oxo-acid-lyase; tryptophan biosynthesis

Query Match 52.9%; Score 45.5; DB 2; Length 505;
 Best Local Similarity 28.0%; Pred. No. 51;
 Matches 7; Conservative 2; Mismatches 3; Indels 13; Gaps 1;

QY 1 ILRW-----PMPMPRRK 12
 1 11111111
 DB 467 VLEWETLNKRRAVGSAMPMPRR 491

RESULT 14

S55483

modulator of drug activity homolog - fission yeast (*Schizosaccharomyces pombe*)C:Species: *Schizosaccharomyces pombe*

C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 10-Dec-1999

C:Accession: S55483; T38969

R:Conor, R.; Church, C.M.

submitted to the EMBL Data Library, May 1995

A:Reference number: S55479

A:Accession: S55483

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-196 <CON>

A:Cross-references: EMBL:249811; NID:9854599; PID:CAA89955.1; PID:9854604

R:Conor, R.; Church, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.

submitted to the EMBL Data Library, May 1995

A:Reference number: Z21821

A:Accession: T38969

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-196 <CO2>

A:Cross-references: EMBL:249811; PID:CAA89955.1; GSPDB:GN00066; SPDB:SPAC5H10.05c

A:Experimental source: strain 972h-; cosmid c5H10

A:Genetics: SPDB:SPAC5H10.05c

A:Gene: SPDB:SPAC5H10.05c

A:Map position: 1

C:Superfamily: NAD(P)H dehydrogenase (quinone) 2

Query Match

Best Local Similarity 52.38; Score 45; DB 2; Length 196;

Matches 8; Conservative 2; Mismatches 2; Indels 4; Gaps 2;

Oy 1 ILRMP-MW---PMRRK 12

Db 63 IYQPMGMWMTPMK 78

RESULT 15

F82646

monofunctional biosynthetic peptidoglycan transglycosylase XF1715 [imported] - *Xylella*C:Species: *Xylella fastidiosa*

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: F82646

R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: F82646

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-273 <SIM>

A:Cross-references: GB:AE003995; GB:AE003849; NID:99106775; PID:AAF84524.1; GSPDB:GN001

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H

as-Neto, E.; Docena, C.; El-Dorry, H.; Fachinani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt

Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, H

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;

F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

A:Genetics:

A:Gene: XF1715

Query Match

Best Local Similarity 52.38; Score 45; DB 2; Length 273;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3 RMPWM 7

Db 44 RMPWM 48

Search completed: June 21, 2001, 08:33:58
Job time: 160 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:37:06 ; Search time 11.81 Seconds

(Without alignments)
34.807 Million cell updates/sec

Title: US-09-444-281-36

Perfect score: 86

Sequence: 1 ILRPMWMPWRK 12

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 93435 segs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	70	81.4	144	1	INDC_BOVIN
2	53	61.6	1173	1	VGL2_CVH22
3	49	57.0	492	1	ADRO_BOVIN
4	47	54.7	253	1	V945_MYCTU
5	47	54.7	713	1	YD53_MYCTU
6	45.5	52.9	505	1	TRPE_PSESS
7	45	52.3	196	1	YA05_SCHPO
8	45	52.3	1108	1	CN3B_RAT
9	44	51.2	361	1	FUT3_HUMAN
10	44	51.2	372	1	FUT3_PANTR
11	44	51.2	397	1	MYL6_MYCTU
12	44	51.2	535	1	YDM6_SCHPO
13	44	51.2	967	1	MYL4_MYCTU
14	44	51.2	968	1	MYL2_MYCTU
15	44	51.2	984	1	MYL2_MYCTU
16	44	51.2	1154	1	VGL2_IBVD2
17	44	51.2	1162	1	VGL2_IBVD2
18	44	51.2	1162	1	VGL2_IBVD2
19	44	51.2	1162	1	VGL2_IBVD2
20	44	51.2	1163	1	VGL2_IBVD2
21	43.5	50.6	276	1	ROEL_RHOPA
22	43	50.0	51	1	LH22_ECTHA
23	43	50.0	711	1	MYL4_STRCO
24	43	50.0	958	1	MYL4_MYCTU
25	43	50.0	1112	1	CN3B_HUMAN
26	43	50.0	1225	1	VGL2_CVPR8
27	43	50.0	1225	1	VGL2_CVPR8
28	43	50.0	1235	1	VGL2_CVPR8
29	43	50.0	1324	1	VGL2_CVPR8
30	43	50.0	1353	1	VGL2_CVPR8
31	43	50.0	1363	1	VGL2_CVPR8
32	43	50.0	1363	1	VGL2_CVPR8
33	43	50.0	1363	1	VGL2_CVPR8

ALIGNMENTS

RESULT	ID	INDC_BOVIN	STANDARD	PRT	144 AA
34	43	50.0	1363	1	VGL2_CVPR8
35	43	50.0	1363	1	VGL2_CVPR8
36	43	50.0	1363	1	VGL2_CVPR8
37	43	50.0	1376	1	VGL2_CVPR8
38	43	50.0	1376	1	VGL2_CVPR8
39	43	50.0	1447	1	VGL2_CVPR8
40	43	50.0	1447	1	VGL2_CVPR8
41	43	50.0	1447	1	VGL2_CVPR8
42	43	50.0	1449	1	VGL2_CVPR8
43	43	50.0	1449	1	VGL2_CVPR8
44	43	50.0	1451	1	VGL2_CVPR8
45	43	50.0	1452	1	VGL2_CVPR8

P15777 bovine coro
 P25193 bovine coro
 P25194 bovine coro
 P22432 bovine coro
 P02385 murine coro
 P07946 porcine tra
 P07946 porcine tra
 P07946 porcine tra
 P18450 porcine tra
 P33470 porcine tra
 P36300 canine ente
 P10033 feline infe

FT DISULFID 85 96 BY SIMILARITY.
 FT DISULFID 107 124 BY SIMILARITY.
 FT MOD RES 143 143 AMIDATION (G-144 PROVIDE AMIDE GROUP).
 SQ. SEQUENCE 144 AA; 16479 MW; E3B1CB8B55C09911 CRC64;

Query Match
 Best Local Similarity 81.4%; Score 70; DB 1; Length 144;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 KMPWMPMR 11
 DB 135 KMPWMPMR 143

RESULT 2
 VGL2_CVH22 STANDARD; PRT; 1173 AA.

AC P15423;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE E2 GLYCOPROTEIN PRECURSOR (SPIKE GLYCOPROTEIN) (PEPLIMER PROTEIN).
 GN S.
 OS Human coronavirus (strain 229E).
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 OC Coronaviridae; Coronavirus.
 NCBI_TaxID=11137;

SEQUENCE FROM N.A.
 MEDLINE=90264837; PubMed=2345367;
 Raabe T., Schelle-prinz B., Siddell S.G.;
 "Nucleotide sequence of the gene encoding the spike glycoprotein of
 human coronavirus HCV 229E".
 J. Gen. Virol. 71:1065-1073(1990).
 J. FUNCTION: THE PEPLIMER PROTEIN MEDIATES THE BINDING OF VIRIONS
 TO THE HOST CELL RECEPTOR AND IS INVOLVED IN MEMBRANE FUSION
 AND IN SYNCTIUM FORMATION.
 CC AND IN SYNCTIUM FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@sib-sib.ch.

DR EMBL: X16816; CAA34723.1;
 DR PIR: A34766; VGIHHC.
 DR InterPro: IPR002551;
 DR Pfam: PF01600; Corona_S1; 1.
 DR Pfam: PF01601; Corona_S2; 1.
 KW Glycoprotein; Envelope protein; Transmembrane; Signal.
 FT SIGNAL 1 15
 FT CHAIN 16 1173
 FT DOMAIN 16 1135
 FT TRANSMEM 1136 1173
 FT DOMAIN 1136 1173
 FT DOMAIN 1136 1157
 FT CARBOHYD 23 23
 FT CARBOHYD 62 62
 FT CARBOHYD 98 98
 FT CARBOHYD 147 147
 FT CARBOHYD 171 171
 FT CARBOHYD 171 171
 FT CARBOHYD 176 176
 FT CARBOHYD 220 220
 FT CARBOHYD 243 243
 FT CARBOHYD 325 325
 FT CARBOHYD 333 333
 FT CARBOHYD 440 440
 FT CARBOHYD 464 464
 E2 GLYCOPROTEIN.
 EXTRACELLULAR (POTENTIAL).
 POTENTIAL.
 CYTOPLASMIC (POTENTIAL).
 CYS-RICH.
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 538 538 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 568 568 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 581 581 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 587 587 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 663 663 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 930 930 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1015 1015 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1020 1020 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1037 1037 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1066 1066 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1076 1076 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1082 1082 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1173 AA; 128639 MW; B9CA9A1A796B3BD CRC64;

Query Match
 Best Local Similarity 61.6%; Score 53; DB 1; Length 1173;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKMPWMPW 9
 DB 1112 LKMPWMPW 1119

RESULT 3
 ADRO_BOVIN STANDARD; PRT; 492 AA.

AC P08165;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE NADPH:ADRENODOXIN OXIDOREDUCTASE, MITOCHONDRIAL PRECURSOR
 DE (C.1.18.1.2) (ADRENODOXIN REDUCTASE) (AR) (FERREDOXIN-NADP(+)
 DE REDUCTASE).
 GN PDXR OR ADXR.
 GN Bos taurus (Bovine).
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;

SEQUENCE FROM N.A. AND ALTERNATIVE SPLICING.
 MEDLINE=94177140; PubMed=8130767;
 Takata Y., Sagara Y., Kono A., Sekimizu K., Horiuchi T.;
 "Gene structure of bovine adrenodoxin reductase.";
 Biol. Pharm. Bull. 16:1200-1206(1993).
 [2]
 SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 MEDLINE=88198050; PubMed=3448086;
 Sagara Y., Takata Y., Miyata T., Hara T., Horiuchi T.;
 "Cloning and sequence analysis of adrenodoxin reductase cDNA from
 bovine adrenal cortex.";
 J. Biochem. 102:1333-1336(1987).
 [3]
 SEQUENCE FROM N.A.
 MEDLINE=87270696; PubMed=3038094;
 Nonaka Y., Murakami H., Yabusaki Y., Kuramitsu S., Kagamiyama H.,
 Yamano T., Okamoto M.;
 "Molecular cloning and sequence analysis of full-length cDNA for mRNA
 of adrenodoxin oxidoreductase from bovine adrenal cortex.";
 Biochem. Biophys. Res. Commun. 145:1239-1247(1987).
 [4]
 SEQUENCE FROM N.A.
 TISSUE-Adrenal cortex;
 MEDLINE=89170752; PubMed=2924777;
 Hanukoglu I., Gutfinger T.;
 "cDNA sequence of adrenodoxin reductase. Identification of NADP-

RT binding sites in oxidoreductases.";
 RL Eur. J. Biochem. 180:479-484(1989).
 RN [5]
 RP SEQUENCE OF N-TERMINUS, AND PARTIAL SEQUENCE.
 RC TISSUE-Adrenal cortex;
 RA MEDLINE-88082777; PubMed-3691502;
 RT Hanukoglu I., Gutfinger T., Hanlu M., Shively J.E.;
 RT "Isolation of a cDNA for adrenodoxin reductase (ferredoxin-NADP+
 RL reductase). Implications for mitochondrial cytochrome P-450 systems.";
 RL Eur. J. Biochem. 169:449-455(1987).
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 33-492.
 RC TISSUE-Adrenal gland;
 RX MEDLINE-99299392; PubMed-10369776;
 RA Ziegler G.A., Vornheim C., Hanukoglu I., Schulz G.E.;
 RT "The structure of adrenodoxin reductase of mitochondrial P450 systems:
 RT electron transfer for steroid biosynthesis.";
 RL J. Mol. Biol. 289:981-990(1999).
 CC -1- FUNCTION: SERVES AS THE FIRST ELECTRON TRANSFER PROTEIN IN ALL THE
 CC MITOCHONDRIAL P450 SYSTEMS, INCLUDING CHOLESTEROL SIDE CHAIN
 CC CLEAVAGE IN ALL STEROIDOGENIC TISSUES, STEROID 11-BETA
 CC HYDROXYLATION IN THE ADRENAL CORTEX, 25-OH-VITAMIN D3-24
 CC HYDROXYLATION IN THE KIDNEY, AND STEROL C-27 HYDROXYLATION IN THE
 CC LIVER.
 CC -1- CATALYTIC ACTIVITY: REDUCED ADRENODOXIN + NADP(+) = OXIDIZED
 CC ADRENODOXIN + NADPH.
 CC -1- COFACTOR: FAD.
 CC -1- PATHWAY: CHOLESTEROL SIDE-CHAIN-CLEAVAGE SYSTEM.
 CC -1- SUBUNIT: MONOMER.
 CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A SHORT FORM (SHOWN HERE) AND A
 CC LONG FORM, ARE PRODUCED BY ALTERNATIVE SPLICING. THE LONG FORM
 CC REPRESENTS 10-20% OF ALL ADRENODOXIN REDUCTASE MRNA. AND SEEMS TO
 CC BE INACTIVE.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: M17029; AAA30362.1;
 DR EMBL: D00211; BAA00150.1;
 DR EMBL: X13736; CAA32002.1;
 DR PIR: A29604; A29604.
 DR PIR: J50390; J50390.
 DR PIR: S03558; S03558.
 DR PIR: J70751; J70751.
 DR PDB: 1CJC; 12-APR-99.
 DR PDB: 1E1L; 02-JUN-00.
 DR InterPro: IPR000759;
 DR PRINTS: PR00419; ADXRDPASE.
 DR Electron transport; oxidoreductase; Flavoprotein; NADP; FAD;
 KW Mitochondrion; Transit peptide; Alternative splicing; 3D-structure.
 FT TRANSIT 1 32
 FT CHAIN 1 32
 FT VARSPIC 33 492
 FT CONFLICT 204 204
 FT CONFLICT 77 77
 FT CONFLICT 81 94
 FT CONFLICT 124 128
 FT CONFLICT 268 268
 FT CONFLICT 317 318
 FT CONFLICT 323 333
 FT CONFLICT 341 352
 SO SEQUENCE 492 AA; 54338 MW; E68F6F5D18F53131 CRC64;

Query Match 57.0% Score 49; DB 1; Length 492;
 Best Local Similarity 83.3% Pred. No. 6;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 WPMWPM 9
 ID 1
 Db 6 WPMWPM 11
 RESULT 4
 Y945_MYCTU
 ID Y945_MYCTU STANDARD; PRT; 253 AA.
 AC P71564;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE OXIDOREDUCTASE RV0945 (EC 1.1.1.1).
 OS RV0945 OR MTCY10D7.29C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL: Z79700; CAB02005.1;
 DR TuberculList; RV0945;
 DR InterPro: IPR002198;
 DR Pfam: PF00106; adh_short; 1.
 DR PROSITE: PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Oxidoreductase.
 FT ACT_SITE 159 159
 FT BY SIMILARITY.
 SO SEQUENCE 253 AA; 27138 MW; BAD937208842DA12 CRC64;

Query Match 54.7% Score 47; DB 1; Length 253;
 Best Local Similarity 100.0% Pred. No. 6;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PMWPM 9
 ID 1
 Db 230 PMWPM 234
 RESULT 5
 YD55_MYCTU
 ID YD55_MYCTU STANDARD; PRT; 715 AA.
 AC Q11025;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 78.2 KDA PROTEIN RV1355C.
 DE RV1355C OR MTCY02B10.19C.

OS Mycobacterium tuberculosis.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CC NCBI_TaxID=1773;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=H37RV;
 CC MEDLINE=98295987; PubMed=9634230;
 CC RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 CC Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 CC Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 CC Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
 CC Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 CC Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 CC Rafter S., Seeger K., Skellern S., Squares S., Squares R., Sulston J.E.,
 CC Taylor K., Whitehead S., Barrett B.G.;
 CC RA Taylor K., Whitehead S., Barrett B.G.;
 CC RT Deciphering the biology of Mycobacterium tuberculosis from the
 CC RT complete genome sequence.;
 CC RT Nature 393:537-544(1998).
 CC RL Nature 393:537-544(1998).
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 CC DR EMBL: Z75555; CAA99988.1;
 CC DR Tuberculist: RV1355C;
 CC DR InterPro: IPR000594;
 CC DR Pfam: PF00899; Thif_family; 1.
 CC DR Hypothetical protein.
 CC KM SEQUENCE 715 AA; 78181 MW; 455495248A56041C CRC64;
 CC SQ
 CC
 CC Query Match 54.7%; Score 47; DB 1; Length 715;
 CC Best Local Similarity 66.7%; Pred. No. 15;
 CC Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC OY 3 RMPMPWRR 11
 CC |:::|||||
 CC DB 65 RMAIYPMRR 73
 CC
 CC RESULT 6
 CC TRPE_PSESS STANDARD; PRT; 505 AA.
 CC ID TRPE_PSESS
 CC AC P21689;
 CC DT 01-MAY-1991 (Rel. 18, Created)
 CC DT 01-MAY-1991 (Rel. 18, Last sequence update)
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)
 CC DE ANTHRANILATE SYNTHASE COMPONENT I (EC 4.1.3.27).
 CC GN TRPE.
 CC OS Pseudomonas syringae (pv. savastanoi).
 CC OC Bacteria; Proteobacteria; gamma subdivision; pseudomonadaceae;
 CC OC Pseudomonas.
 CC ON NCBI_TaxID=29438;
 CC RX [1]
 CC RP SEQUENCE FROM N.A.
 CC MEDLINE=91100331; PubMed=1987141;
 CC RA da Costa E., Silva O., Kosuge T.;
 CC RT Molecular characterization and expression analysis of the
 CC RT anthranilate synthase gene of Pseudomonas syringae subsp.
 CC RT savastanoi.;
 CC RT J. Bacteriol. 173:463-471(1991).
 CC CC
 CC CC -1- CATALYTIC ACTIVITY: CHORISMATE + L-GLUTAMINE -> ANTHRANILATE +
 CC PYRUVATE + L-GLUTAMATE.
 CC CC -1- PATHWAY: FIRST STEP IN BIOSYNTHESIS OF TRYPTOPHAN.
 CC CC -1- SUBUNIT: Tetramer of two components I and two components II (by
 CC CC SIMILARITY).
 CC CC -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE
 CC USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES

CC GLUTAMINE AMIDOPRANSEASE ACTIVITY.
 CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I
 CC FAMILY.
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 CC DR EMBL: M55911; AAA26016.1;
 CC DR PIR: A39128; A39128.
 CC DR InterPro: IPR000350;
 CC DR Pfam: PF00425; Chorismate_bind; 1.
 CC DR PRINTS: PR00095; ANTSYNTHASE1.
 CC KW Tryptophan biosynthesis; Lyase.
 CC KM SEQUENCE 505 AA; 56084 MW; A38E8193131F6BB CRC64;
 CC SQ

Query Match 52.9%; Score 45.5; DB 1; Length 505;
 Best Local Similarity 28.0%; Pred. No. 18;
 Matches 7; Conservative 2; Mismatches 3; Indels 13; Gaps 1;
 OY 1 ILRW-----PMPWRRK 12
 |:::|||||
 DB 467 VLEWETLNKRRAMVGSAMPWRR 491

RESULT 7
 YA05_SCHPO STANDARD; PRT; 196 AA.
 ID YA05_SCHPO
 AC Q09677;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE HYPOTHELICAL 22.1 KDA PROTEIN C5H10.05C IN CHROMOSOME I.
 GN SPAC5H10.05C.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC Schizosaccharomycetes.
 CC ON NCBI_TaxID=4896;
 CC RX [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RA Connor R., Churcher C.M., Barrett B.G., Rajandream M.A., Walsh S.V.;
 CC Submitted (May-1995) to the EMBL/Genbank/DBJ databases.
 CC RL -1- SIMILARITY: STRONG, TO BACTERIAL MODULATOR OF DRUG ACTIVITY B
 CC (MDAB).
 CC CC
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC CC
 CC DR EMBL: Z49811; CAA89955.1;
 CC DR Hypothetical protein.
 CC KW SEQUENCE 196 AA; 22104 MW; 436764DA9E26074C CRC64;
 CC SQ

Query Match 52.3%; Score 45; DB 1; Length 196;
 Best Local Similarity 50.0%; Pred. No. 8.6;
 Matches 8; Conservative 2; Mismatches 2; Indels 4; Gaps 2;
 OY 1 ILRW-PW---PMPRRK 12
 |:::|||||
 DB 63 IYMPGMMGTMPRLK 78

```

RESULT 8
CN3B_RAT 9
ID CN3B_RAT STANDARD: PRT: 1108 AA.
AC 063085;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE CGMP-INHIBITED 3',5'-CYCLIC PHOSPHODIESTERASE B (EC 3.1.4.17) (CYCLIC
DE GMP INHIBITED PHOSPHODIESTERASE B) (CGI-PDE B) (CGIPDE1).
GN PDE3B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=adipose tissue;
RX MEDLINE=93366761; PubMed=8395509;
RA Taira M., Hockman S.C., Calvo J.C., Taira M., Belfrage P.,
RA Manganelli V.C.;
RT "Molecular cloning of the rat adipocyte hormone-sensitive cyclic GMP-
RT inhibited cyclic nucleotide phosphodiesterase."
RL J. Biol. Chem. 268:18573-18579(1993).
CC -1- FUNCTION: MAY PLAY A ROLE IN FAT METABOLISM.
CC -1- CATALYTIC ACTIVITY: GUANOSINE 3',5'-CYCLIC PHOSPHATE + H(2)O =
CC GUANOSINE 5'-PHOSPHATE.
CC -1- ENZYME REGULATION: INHIBITED BY GMP.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND (POTENTIAL).
CC -1- TISSUE SPECIFICITY: ABUNDANT IN ADIPOSE TISSUES.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE PHOSPHODIESTERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: 222867; CAA80489.1; -
DR InterPro: IPR002073; -
DR Pfam: PF00233; PDEase_1;
DR PROSITE: PS00126; PDEASE_1; 1.
KM Hydrolyase; GMP; Membrane.
FT DOMAIN 16 22 POLY-PRO.
FT DOMAIN 99 102 POLY-ALA.
FT DOMAIN 175 179 POLY-ALA.
FT DOMAIN 1007 1021 POLY-ASP.
FT DOMAIN 1068 1071 POLY-GLU.
FT DOMAIN 1101 1104 POLY-GLU.
SO SEQUENCE 1108 AA; 123105 MW; C9B5078C7D3ADD6D CRC64;

```

```

Query Match 52.3%; Score 45; DB 1; Length 1108;
Best Local Similarity 62.5%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 4 WPMWPMRR 11
Db 164 WPMWPMRR 171

```

```

RESULT 9
FUT3_HUMAN STANDARD: PRT: 361 AA.
ID FUT3_HUMAN
AC P21217; Q99448; Q99449;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GALACTOSIDE 3(4)-L-FUCOSYLTRANSFERASE (EC 2.4.1.65) (BLOOD GROUP LEWIS
DE ALPHA-4-FUCOSYLTRANSFERASE) (LEWIS FT) (FUCOSYLTRANSFERASE 3) (FUCT-

```

```

DE III).
GN FUT3 OR LE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=94059067; PubMed=7650030;
RA Cameron H.S., Szczepaniak D., Weston W.;
RT "Expression of human chromosome 19p alpha(1,3)-fucosyltransferase
RT genes in normal tissues. Alternative splicing, polyadenylation, and
RT isoforms."
RL J. Biol. Chem. 270:20112-20122(1995).
RN [3]
RN VARIANT LE(-) MET-105.
RX MEDLINE=94059067; PubMed=8240322;
RA Elmgren A., Rydberg L., Larson G.;
RT "Genotypic heterogeneity among Lewis negative individuals."
RL Biochem. Biophys. Res. Commun. 196:515-520(1993).
RN [4]
RN VARIANT LE(-) ARG-20; SER-170 AND ALA-336.
RX MEDLINE=94059082; PubMed=8240337;
RA Nishihara S., Yazawa S., Iwasaki H., Nakazato M., Kudo T., Ando T.,
RA Nishimatsu H.;
RT "Alpha (1,3/1,4)-fucosyltransferase (FUCT-III) gene is inactivated by
RT a single amino acid substitution in Lewis histo-blood type negative
RT individuals."
RL Biochem. Biophys. Res. Commun. 196:624-631(1993).
RN [5]
RN VARIANT LE(-) ARG-20 AND SER-170.
RX MEDLINE=94033579; PubMed=8219240;
RA Koda Y., Kimura H., Mekada E.;
RT "Analysis of Lewis fucosyltransferase genes from the human gastric
RT mucosa of Lewis-positive and -negative individuals."
RL Blood 82:2915-2919(1993).
RN [6]
RN VARIANT LE(-) ARG-20 AND LYS-356.
RX MEDLINE=94342259; PubMed=8063716;
RA Mollicone R., Reguigne I., Kelly R.J., Fletcher A., Watt J.,
RA Chatfield S., Aziz A., Cameron H.S., Weston B.W., Lowe J.B., Oriol R.;
RT "Molecular basis for Lewis alpha(1,3/1,4)-fucosyltransferase gene
RT deficiency (FUT3) found in Lewis-negative Indonesian pedigrees."
RL J. Biol. Chem. 269:20987-20994(1994).
RN [7]
RN VARIANT LE(-) LYS-356.
RX MEDLINE=95050753; PubMed=7961897;
RA Nishihara S., Nishimatsu H., Iwasaki H., Yazawa S., Akamatsu S.,
RA Ando T., Seno T., Nishimatsu I.;
RT "Molecular genetic analysis of the human Lewis histo-blood group
RT system."
RL J. Biol. Chem. 269:29271-29278(1994).
RN [8]
RN VARIANT LE(-) ARG-20; ARG-68; MET-105 AND LYS-356.
RX MEDLINE=96243526; PubMed=8801770;
RA Elmgren A., Boerjeson C., Svensson L., Rydberg L., Larson G.;
RT "DNA sequencing and screening for point mutations in the human Lewis
RT 'FUT3' gene enables molecular genotyping of the human Lewis blood
RT group system."
RL Vox Sang. 70:97-103(1996).
RN [9]
RN VARIANT LE(-) ARG-68 AND MET-105.
RX MEDLINE=97413801; PubMed=9268337;
RA Elmgren A., Mollicone R., Costache M., Boerjeson C., Oriol R.,

```

RA	HA	Harrington J., Larson G.:	*Significance of individual point mutations, 7202C and C314T, in the
RT	human	Lewis 'Fut3' gene for expression of Lewis antigens by the human	
RT	alpha1,3/1,4-fucosyltransferase, Fuc-III.1."		
RL	J. Biol. Chem.	272:21994-21998(1997).	
RP	1101		
RA	VAR1ANTS LE(+)* R-102; A-124, AND VAR1ANTS LE(-)* N-162; R-223; M-270.		
RX	MEDLINE=9636689; PubMed=9703429;		
RA	Pang H., Liu Y., Koda Y., Soejima M., Jia J., Schlaphoff T.,		
RT	"Five novel missense mutations of the Lewis gene 'Fut3' in African		
RT	'Xhosa' and Caucasian populations in South Africa."		
RL	Hum. Genet. 102:675-680(1998).		
CC	-1- FUNCTION: MAY CATALYZE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES		
CC	INVOLVED IN THE EXPRESSION OF VIN-2, LEWIS A, LEWIS B, SIATAL		
CC	LEWIS X AND LEWIS X/SSSEA-1 ANTIGENS. MAY BE INVOLVED IN BLOOD		
CC	GROUP LEWIS DETERMINATION: LEWIS-POSITIVE (LE(+)) INDIVIDUALS		
CC	HAVE AN ACTIVE ENZYME WHILE LEWIS-NEGATIVE (LE(-)) INDIVIDUALS		
CC	HAVE AN INACTIVE ENZYME.		
CC	-1- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,3-BETA-D-GALACTOSYL-		
CC	-N-ACETYL-D-GLUCOSAMINYL-R = GDP + 1,3-BETA-D-GALACTOSYL-		
CC	(ALPHA-1,4-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.		
CC	-1- PATHWAY: GLYCOSYLATION.		
CC	-1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND		
CC	FORM IN TRANS CISTERNAE OF GOLGI.		
CC	-1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN STOMACH, COLON, SMALL		
CC	INTESTINE, LUNG AND KIDNEY AND TO A LESSER EXTENT IN SALIVARY		
CC	GLAND, BLADDER, UTERUS AND LIVER.		
CC	-1- MISCELLANEOUS: ALSO ACTS ON THE CORRESPONDING 1,4-GALACTOSYL		
CC	DERIVATIVE, FORMING 1,3-L-FUCOSYL LINKS.		
CC	-1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN		
CC	GLYCOSYLTRANSFERASES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; X53578; CAA37641.1; -		
DR	EMBL; U27328; AAC50187.1; -		
DR	EMBL; U27326; AAC50185.1; -		
DR	EMBL; U27327; AAC50186.1; -		
DR	EMBL; D89324; BAA13941.1; -		
DR	EMBL; D89325; BAA13942.1; -		
DR	PIR; A36669; A36669.		
DR	MIM; 111100; -		
DR	InterPro; IPR001503; -		
DR	Pfam; PF00852; Fucosyl_1_transf. 1.		
RW	Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;		
KV	Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.		
FT	DOMAIN	1 15	
FT	TRANSHEM	16 34	
FT	DOMAIN	35 361	
FT	CARBOHYD	154 154	
FT	CARBOHYD	185 185	
FT	VARIANT	20 20	
FT	VARIANT	68 68	
FT	VARIANT	102 102	
FT	VARIANT	105 105	
FT	VARIANT	124 124	
FT	VARIANT	162 162	
FT	VARIANT	170 170	
FT	VARIANT	223 223	
FT	G-> R (IN LE(-)).		
FT	G-> S (IN LE(-)). COMPLETELY INACTIVE).		
FT	/Fut3g-VAR_003428.		
FT	G-> R (IN LE(-)).		
FT	W-> R (IN LE(-)).		
FT	/Fut3g-VAR_007959.		
FT	Q-> R (IN LE(+)).		
FT	/Fut3g-VAR_007960.		
FT	T-> M (IN LE(-)).		
FT	/Fut3g-VAR_003427.		
FT	S-> A (IN LE(+)).		
FT	/Fut3g-VAR_007961.		
FT	D-> N (IN LE(-)).		
FT	/Fut3g-VAR_007962.		
FT	G-> S (IN LE(-)). COMPLETELY INACTIVE).		
FT	/Fut3g-VAR_003428.		
FT	G-> R (IN LE(-)).		

```

FT FT VARIANT 270 270 /FTID-VAR_007963.
FT FT VARIANT 336 336 /FTID-VAR_007964.
FT FT VARIANT 356 356 D -> A (IN LE(-)).
FT FT VARIANT 356 356 /FTID-VAR_003429.
FT FT I -> K (IN LE(-)); LESS THAN 10% REDUCTION
FT FT IN ACTIVITY.
SQ SEQUENCE 361 AA; 42117 MW; BF4398044F19C284 CRC64;

Query Match 51.2%; Score 44; DB 1; Length 361;
Best Local Similarity 85.7%; Pred No. 20;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0

QY 5 PWMPMR 11
   1 1111
Db 9 PWMPMR 15

RESULT 10
FUT3_PANTR STANDARD; PRT; 372 AA.
ID FUT3_PANTR 019058;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE GALACTOSIDE 3(4)-L-FUCOSYLTRANSFERASE (EC 2.4.1.65) (BLOOD GROUP LEWIS
DE ALPHA-4-FUCOSYLTRANSFERASE) (LEWIS FU) (FUCOSYLTRANSFERASE 3) (FUCT-
DE III) (ALPHA-3/4-FUCOSYLTRANSFERASE).
GN FUT3.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_Taxid=9598;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96037800; PubMed=9368041;
RA Costache M., Apoll P.-A., Calileau A., Elmgren A., Larson G.,
RA Henry S., Blancher A., Iordachescu D., Ortol R., Mollitico R.;
RT "Evolution of fucosyltransferase genes in vertebrates.";
RL J. Biol. Chem. 272:29721-29728(1997)
CC -1- FUNCTION: MAY CATALYZE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES
CC INVOLVED IN THE EXPRESSION OF SIALYL LEWIS X AND LEWIS X/SSA-1
CC ANTIGENS. IT MAY BE INVOLVED IN BLOOD GROUP LEWIS DETERMINATION
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,3-BETA-D-GALACTOSYL-
CC N-ACETYL-D-GLUCOSAMINYL-R -> GDP + 1,3-BETA-D-GALACTOSYL-
CC (ALPHA-1,4-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN, MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).
CC -1- POLYMORPHISM: THERE ARE TWO ALLELES (A AND B). ALLELE A HAS ARG-
CC 162 AND VAL-304. ALLELE B HAS GLY-162 AND MET-304.
CC -1- MISCELLANEOUS: ALSO ACTS ON THE CORRESPONDING 1,4-GALACTOSYL
CC DERIVATIVE, FORMING 1,3-L-FUCOSYL LINKS.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
CC GLYCOSYLTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y14033; CAAT4360.1; -
DR InterPro; IPR001503; -
DR Pfam; PF00882; Fucosyl_transf_1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Polymorphism.
KW DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).

```

FT TRANSMEM 15 34 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 (POTENTIAL).
 FT DOMAIN 35 372 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 162 162 R -> G (IN ALLELE B).
 FT VARIANT 304 304 V -> M (IN ALLELE B).
 SQ SEQUENCE 372 AA: 43233 MW: 6490BF8BCA7BD74C CRC64:

Query Match 51.2%: Score 44: DB 1: Length 372;
 Best Local Similarity 85.7%: Pred No. 21;
 Matches 6: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Oy 5 PMPWRR 11
 1 11111
 Db 9 PQMPWR 15

RESULT 11
 MML6_MYCTU STANDARD: PRT: 397 AA.
 AC Q10773:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN Mmpl6.
 GN Mmpl6 OR RV1557 OR MTCY48.08C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=96295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Horsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Sultson J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE Mmpl FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: Z74020; GCA98334.1;
 KM Hypothetical protein, Transmembrane.
 FT TRANSMEM 161 181 POTENTIAL.
 FT TRANSMEM 190 210 POTENTIAL.
 FT TRANSMEM 214 234 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 293 313 POTENTIAL.
 FT TRANSMEM 330 350 POTENTIAL.
 SQ SEQUENCE 397 AA: 42421 MW: 678DC86E24472BF4 CRC64:

Query Match 51.2%: Score 44: DB 1: Length 397;
 Best Local Similarity 75.0%: Pred No. 22;
 Matches 6: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

Oy 3 RMPWRR 10
 1 11111
 Db 351 RMPWRR 358

RESULT 12
 YDM6_SCHPO STANDARD: PRT: 535 AA.
 ID YDM6_SCHPO
 AC 013912;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL 60.1 KDA PROTEIN C23C11.06C IN CHROMOSOME I.
 GN SPAC23C11.06C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Brown D., Churcher C.M., Barrett B.G., Rajandream M.A., Wood V.;
 RC STRAIN=972;
 RL Submitted (AUG-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

DR EMBL: Z98559; CAB1159.1;
 KM Hypothetical protein, Transmembrane.
 FT TRANSMEM 55 75 POTENTIAL.
 FT TRANSMEM 82 102 POTENTIAL.
 FT TRANSMEM 115 135 POTENTIAL.
 FT TRANSMEM 143 163 POTENTIAL.
 FT TRANSMEM 201 221 POTENTIAL.
 FT TRANSMEM 346 366 POTENTIAL.
 SQ SEQUENCE 535 AA: 60124 MW: A6AE149AA2929E2 CRC64:

Query Match 51.2%: Score 44: DB 1: Length 535;
 Best Local Similarity 50.0%: Pred No. 29;
 Matches 6: Conservative 1: Mismatches 1: Indels 4: Gaps 1;

Oy 4 WPMWRR 11
 1 11111
 Db 183 WPMWRR 194

RESULT 13
 MML4_MYCTU STANDARD: PRT: 967 AA.
 AC 053735;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN Mmpl4.
 GN Mmpl4 OR RV0450C OR MTCY37.14C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.

RA MEDLINE=96295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE MPTL FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AL017932; CA17407.1; -
 DR Tuberculist; RV0450C; -
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 26 46 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 242 262 POTENTIAL.
 FT TRANSMEM 303 323 POTENTIAL.
 FT TRANSMEM 333 353 POTENTIAL.
 FT TRANSMEM 384 404 POTENTIAL.
 FT TRANSMEM 769 789 POTENTIAL.
 FT TRANSMEM 793 813 POTENTIAL.
 FT TRANSMEM 821 841 POTENTIAL.
 FT TRANSMEM 875 895 POTENTIAL.
 FT TRANSMEM 896 916 POTENTIAL.
 SQ SEQUENCE 967 AA; 105234 MW; 6301014031480484 CRC64;

Query Match 51.2%; Score 44; DB 1; Length 967;
 Best Local Similarity 75.0%; Pred. No. 50;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RMPMPWR 10
 Db 930 RMPMPWR 937

RESULT 14
 MML2_MYCTU STANDARD; PRT; 968 AA.
 ID MML2_MYCTU
 AC 011171;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PUTATIVE MEMBRANE PROTEIN MML2.
 GN MML2 OR RV0507 OR MYCY20G.34.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Baddock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.,
 RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE MPTL FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z77162; CAB00933.1; -
 DR Tuberculist; RV0507; -
 KW Hypothetical protein; Transmembrane.
 FT TRANSMEM 22 42 POTENTIAL.
 FT TRANSMEM 204 224 POTENTIAL.
 FT TRANSMEM 245 265 POTENTIAL.
 FT TRANSMEM 297 317 POTENTIAL.
 FT TRANSMEM 328 348 POTENTIAL.
 FT TRANSMEM 378 398 POTENTIAL.
 FT TRANSMEM 763 783 POTENTIAL.
 FT TRANSMEM 787 807 POTENTIAL.
 FT TRANSMEM 815 835 POTENTIAL.
 FT TRANSMEM 866 886 POTENTIAL.
 FT TRANSMEM 891 911 POTENTIAL.
 SQ SEQUENCE 968 AA; 106201 MW; B68AE9B78164EDC0 CRC64;

Query Match 51.2%; Score 44; DB 1; Length 968;
 Best Local Similarity 75.0%; Pred. No. 50;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 RMPMPWR 10
 Db 924 RMPMPWR 931

RESULT 15
 SX13_MOUSE STANDARD; PRT; 984 AA.
 ID SX13_MOUSE
 AC 004831;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-OCT-2000 (Rel. 40, Last annotation update)
 DE SOX-13 PROTEIN.
 GN SOX13 OR SOX-13.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND FUNCTION (ISOFORM 1).
 RC TISSUE=Embryo;
 RX MEDLINE=98083175; PubMed=9421502;
 RA Roose J., Koryer W., Oving E., Wilson A., Wagenaar G., Markman M.,
 RA Lamers W., Clevers H.;
 RT "High expression of the HMG box factor sox-13 in arterial walls during
 RT embryonic development.";
 RL Nucleic Acids Res. 26:469-476(1998).
 RN [2]
 RP SEQUENCE FROM N.A., AND FUNCTION (ISOFORM 2).
 RC TISSUE=Embryo;
 RX MEDLINE=98201614; PubMed=9524265;
 RA Kido S., Hiraoka Y., Ogawa M., Sakai Y., Yoshimura Y., Aiso S.;
 RT "Cloning and characterization of mouse msox13 cDNA.";
 RL Gene 208:201-206(1998).
 RN [3]
 RP SEQUENCE OF 405-460 FROM N.A.
 RX MEDLINE=93181275; PubMed=8441686;
 RA Wright E.M., Snopce B., Koopman P.;
 RT "Seven new members of the Sox gene family expressed during mouse

```

RT development."
RL Nucleic Acids Res. 21:744-744(1993).
CC -1- FUNCTION: BINDS TO THE SEQUENCE 5'-AACAT-3'.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; ISOFORM 1 (SHOWN HERE) AND
CC ISOFORM 2; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, HIGH LEVELS OF EXPRESSION ARE
CC FOUND IN THE ARTERIAL WALLS AT 13.5 DAYS POST COITUM (DPC). LOW
CC LEVELS ARE FOUND IN THE INNER EAR AT 13.5 DPC AND IN SOME CELLS IN
CC THE THYMUS AT 16.5 DPC. EXPRESSED IN THE TRACHEAL EPITHELIUM BELOW
CC THE VOCAL CORD AND IN THE HAIR FOLLICLES AT 18 DPC.
CC -1- SIMILARITY: CONTAINS 1 HMG BOX.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ000740; CAA04278.1; -
DR EMBL: AB006329; BAA25786.1; -
DR EMBL: Z18962; CAA79487.1; -
DR PIR: S30241; S30241.
DR HSP: Q05066; 1HR2.
DR MGD: MGI:98361; SOX13.
DR InterPro: IPR000910; -
DR Pfam: PF00505; HMG_box; 1.
KW DNA-binding; Nuclear protein; Alternative splicing.
FT DOMAIN 159 195 GLN-RICH.
FT DNA_BIND 397 465 HMG_BOX.
FT VARSPPLIC 495 519 PGCSPKLHPVSRPSLVARGGLMLL -> OGAROSYTIIP
FT (IN ISOFORM 2).
FT VARSPPLIC 603 609 SMWYSQT -> ELVVLTD (IN ISOFORM 2).
FT VARSPPLIC 610 984 MISSING (IN ISOFORM 2).
FT CONFLICT 35 35 P -> L (IN REF. 2).
FT CONFLICT 41 42 AT -> TN (IN REF. 2).
FT CONFLICT 195 195 Q -> QQ (IN REF. 2).
SO SEQUENCE 984 AA; 108897 MW; 7F5506EDADEB98C5 CRC64;

Query Match 51.2%; Score 44; DB 1; Length 984;
Best Local Similarity 42.98; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
OY 4.WPMW-----PM 9
DB 686 WPMWTKLAEGFSPW 699

```

Search completed: June 21, 2001, 08:37:06
Job time: 233 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: June 21, 2001, 08:37:43 ; Search time 33.4 Seconds
(without alignments)
47.535 Million cell updates/sec

```

Title: US-09-444-281-36
Perfect score: 86
Sequence: 1 ILRWPPWPPWRRK 1

```

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 425026 seqs, 132305027 residues
Total number of hits satisfying chosen parameters

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

```
Database :
    sPTREMBL_16:*
1:  sp.archae:*
2:  sp.bacteria:*
3:  sp.fungi:*
4:  sp.human:*
5:  sp.invertebrate:*
6:  sp.mammal:*
7:  sp.mhc:*
8:  sp.organelle:*
9:  sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.unclassified:*
13: sp.vertebrate:*
14: sp.virus:*
```

Pred. No.: is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	57	66.3	723	14	09DUC4	09auc4 tt. virus.
2	53	61.6	746	14	09JH31	09jh31 tt. virus.
3	53	61.6	1383	14	084712	084712 porcine ep
4	52	60.5	1245	5	09Y7V5	09y7v5 trichoderm
5	51	59.3	239	4	09Y4N1	09y4n1 homo sapien
6	51	59.3	504	2	P6143	P6143 thermocactin
7	50	58.1	111	5	018753	018753 caenorhabdi
8	50	58.1	735	14	09DUC9	09duc9 tt. virus.
9	49	57.0	49	14	09DT80	09dt80 tt. virus.
10	49	57.0	467	5	019573	019573 caenorhabdi
11	49	57.0	748	14	09DT81	09dt81 tt. virus.
12	48.5	56.4	114	2	09X8C2	09x8c2 streptomyce
13	48	55.8	540	2	007504	007504 bacillus me
14	47	54.7	165	10	09SNN3	09snn3 oryza satii
15	47	54.7	276	2	09HXK9	09hxc9 pseudomonas
16	47	54.7	1411	10	09LYG0	09lyg0 arabidopsi
17	46	53.5	154	2	09R6J3	09r6j3 agrobacteri
18	46	53.5	728	3	09P3G0	09p3g0 neurospora
19	45	52.3	159	2	09K273	09k273 streptomyce

ALIGNMENTS

20	45	52.3	273	2	09PCR3	09PCR3 xylella fas
21	45	52.3	412	2	09i6f7	09i6f7 pseudomonas
22	45	52.3	423	2	024742	024742 bacteroides
23	45	52.3	443	10	09s751	09s751 oryza sativ
24	45	52.3	448	2	09wYR8	09wYR8 thermotoga
25	45	52.3	730	10	09F6Z6	09F6Z6 arabidopsis
26	45	52.3	767	14	09QUD8	09QUD8 tt virus. h
27	45	52.3	1100	11	09ZLJ9	09ZLJ9 mus musculu
28	44.5	51.7	766	14	09IRV0	09IRV0 tt virus. p
29	44.5	51.2	143	4	09H9A4	09H9A4 homo sapien
30	44	51.2	145	2	086437	086437 pseudomonas
31	44	51.2	257	2	056924	056924 yersinia en
32	44	51.2	361	4	09PIW6	09PIW6 homo sapien
33	44	51.2	406	5	09W404	09W404 drosophila
34	44	51.2	429	5	09N8Y2	09N8Y2 trypanosoma
35	44	51.2	458	4	09UGB1	09UGB1 homo sapien
36	44	51.2	620	14	09IH07	09IH07 avian infec
37	44	51.2	621	14	066196	066196 avian infec
38	44	51.2	621	14	09IH14	09IH14 avian infec
39	44	51.2	621	14	09IH13	09IH13 avian infec
40	44	51.2	621	14	09IH11	09IH11 avian infec
41	44	51.2	621	14	09IH10	09IH10 avian infec
42	44	51.2	621	14	09IH15	09IH15 avian infec
43	44	51.2	621	14	09IH12	09IH12 avian infec
44	44	51.2	621	14	09IH09	09IH09 avian infec
45	44	51.2	621	14	09IH08	09IH08 avian infec

Query Match	66.3%	Score 57;	DB 14;	Length 723;
Best Local Similarity	100.0%	Pred. NO. 4.3;		
Matches	7;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
OY	5	PWMPWR	11	
Db	2	PWMPWR	8	
RESULT	2			
O9JH31				
ID	O9JH31	PRELIMINARY:	PRT:	746 AA.
GC	O9JH31:			

DT 01-OCT-2000 (TEMBLrel. 15, Created)
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)
 DE 01-OCT-2000 (TEMBLrel. 15, Last annotation update)
 DE ORF1.
 OS TT virus.
 CC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TJN02;
 RA Okamoto H.;
 RN Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TJN02;
 RA Urita M., Okamoto H., Nishizawa T., Tawara A., Takahashi M.,
 RA Iizuka H., Miyakawa Y., Mayumi M.;
 RT "The entire nucleotide sequences of two distinct TT virus (TTV)
 RT isolates (TJN01 and TJN02) remotely related to the original TTV
 RT isolates."
 RL Arch. Virol. 0:0-0(2000).
 DR EMBL: AB028669; BAA94878.1;
 SO SEQUENCE 746 AA; 88561 MW; E0B22953AE764E3E CRC64;

Query Match 61.6%; Score 53; DB 14; Length 746;
 Best Local Similarity 54.5%; Pred. No. 14;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 LRMPMPWRRK 12
 DB 1 MAMGWMRRRR 11

RESULT 3
 ID 084712 PRELIMINARY; PRT; 1383 AA.
 AC 084712;
 DT 01-NOV-1996 (TEMBLrel. 01, Created)
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
 DT 01-MAY-2000 (TEMBLrel. 13, Last annotation update)
 DE SPIKE PROTEIN.
 OS Porcine epidemic diarrhea virus.
 CC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
 CC Coronaviridae; Coronavirus.
 OX NCBI_TaxID=28295;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRI/87;
 RX MEDLINE=94231173; PubMed=8176382;
 RA Duarte M., Laude H.;
 RT "Sequence of the spike protein of the porcine epidemic diarrhoea
 RT virus."
 RL J. Gen. Virol. 75:1195-1200(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRI/87;
 RX MEDLINE=93389433; PubMed=8397280;
 RA Bridgen A., Duarte M., Tobler K., Laude H., Ackermann M.;
 RT "Sequence determination of the nucleocapsid protein gene of the
 RT porcine epidemic diarrhoea virus confirms that this virus is a
 RT coronavirus related to human coronavirus 229E and porcine
 RT transmissible gastroenteritis virus."
 RL J. Gen. Virol. 74:1795-1804(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRI/87;
 RX MEDLINE=94120721; PubMed=8291230;
 RA Duarte M., Tobler K., Bridgen A., Rasschaert D., Ackermann M.,
 RA Laude H.;
 RT "Sequence analysis of the porcine epidemic diarrhoea virus genome
 RT between the nucleocapsid and spike protein genes reveals a polymorphic
 RT ORF.";

RL Virology 198:466-476(1994).
 DR EMBL: Z25483; CAA80971.1;
 DR Interpro: IPR002551;
 DR Interpro: IPR002552;
 DR Pfam: PF01600; Corona_S1; 1.
 DR Pfam: PF01601; Corona_S2; 1.
 FT CONFLICT 422 422
 SO SEQUENCE 1383 AA; 151404 MW; 741C84D5DD3BC4D CRC64;

Query Match 61.6%; Score 53; DB 14; Length 1383;
 Best Local Similarity 62.5%; Pred. No. 25;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LRMPMPW 9
 DB 1321 IKMPMPW 1328

RESULT 4
 ID 09Y7V5 PRELIMINARY; PRT; 1245 AA.
 AC 09Y7V5;
 DT 01-NOV-1999 (TEMBLrel. 12, Created)
 DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE CONIDIOSPORE SURFACE PROTEIN.
 GN CMPL.
 OS Trichoderma harzianum.
 CC Eukaryota; Fungi; Ascomycota; mitosporic Ascomycota; Trichoderma.
 OX NCBI_TaxID=5544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ARCC 32173;
 RA Puyesky M., Benhamou N., Ponce Noyola P., Bauw G., Ziv T.,
 RA van Montagu M., Herrera Estrella A., Horwitz B.A.;
 RT "Developmental regulation of a gene encoding a multidomain
 RT conidiospore surface protein of Trichoderma, cmpl."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ133651; CAB40845.1;
 SO SEQUENCE 1245 AA; 135824 MW; 3249C749AFA0CDF8 CRC64;

Query Match 60.5%; Score 52; DB 3; Length 1245;
 Best Local Similarity 60.0%; Pred. No. 30;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 3 RMPMPWRRK 12
 DB 1185 RMQWMPRR 1194

RESULT 5
 ID 09Y4N1 PRELIMINARY; PRT; 299 AA.
 AC 09Y4N1;
 DT 01-NOV-1999 (TEMBLrel. 12, Created)
 DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
 DE HYPOHERICAL 34.0 KDA PROTEIN (FRAGMENT).
 GN DKFZP434C192.
 OS Homo sapiens (human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RA Ansoorge W., Winkler U., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL0096753; CAB46428.2;
 KW Hypoherical protein.
 FT NON_TER 1

SQ SEQUENCE 299 AA: 34032 MW: 688DB60E6A88239A CRC64;
 Query Match 59.3%; Score 51; DB 4; Length 299;
 Best Local Similarity 85.7%; Pred. No. 12;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 5 PMPMPMR 11
 |||||
 Db 37 PMPMPMR 43

RESULT 6
 ID P96143 PRELIMINARY; PRT: 504 AA.
 AC P96143;
 DT 01-MAY-1997 (TREMBlrel. 03, Created)
 DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
 DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
 DE PEPTIDE HYDROLASE.
 GN TLEP1.
 OS Thermactinomyces vulgaris.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Thermactinomyces.
 OX NCBI_TaxID=2026;
 RN [1]
 RN SEQUENCE OF 1-431 FROM N.A.
 RC STRAIN-94-2A;
 RA Hofmeister J.W.;
 RL Institut fuer Pflanzen-genetik und Kulturpflanzenforschung.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-94-2A;
 RA Hofmeister J.W.;
 RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 283314; CAB05671.1;
 DR HSSP: P00800; 1HTT.
 DR InterPro: IPR001570;
 DR Pfam: PF01447; Peptidase_M4; 1.
 KM Hydrolase.
 SQ SEQUENCE 504 AA: 56653 MW: 5A7BCC05C5AD1315 CRC64;

Query Match 59.3%; Score 51; DB 2; Length 504;
 Best Local Similarity 60.0%; Pred. No. 18;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 ILRMPMPMR 10
 ::|||
 Db 71 LVFMTWMPMR 80

RESULT 7
 ID Q18753 PRELIMINARY; PRT: 111 AA.
 AC Q18753;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
 DE GLYCINE-RICH.
 GN C50F7.8.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=94150718; Pubmed=7906398;
 RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
 Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
 Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 Jones M., Kershaw J., Kirsten J., Laister J., Latreille P.,

RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
 Parsons J., Percy C., Rikken L., Roopar A., Saunders D., Showkeen R.,
 Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
 Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson-Spirot J., Wohlman P.,
 RT 2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans.
 RT Nature 368:32-38(1994).
 RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Johnson D., Stelljes L.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RA Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U41557; AAA83303.1;
 SQ SEQUENCE 111 AA: 10139 MW: 66729A2E0F9762B9 CRC64;

Query Match 58.1%; Score 50; DB 5; Length 111;
 Best Local Similarity 54.5%; Pred. No. 6.6;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ILRMPMPMR 11
 ::|||
 Db 12 VMMPMPMPGR 22

RESULT 8
 ID Q9DUC9 PRELIMINARY; PRT: 735 AA.
 AC Q9DUC9;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE ORF1.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PT-TTV6;
 RA Okamoto H.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PT-TTV6;
 RX Pubmed=11080484;
 RA Okamoto H., Nishizawa T., Tawara A., Peng Y., Takahashi M.,
 Kishimoto J., Tanaka T., Miyakawa Y., Mayumi M.;
 RT phylogenetic relatedness.
 RT Virology 277:368-378(2000).
 DR EMBL: AB041957; BAB19308.1;
 SQ SEQUENCE 735 AA: 86132 MW: 9ED81BD6BE6FA5D3 CRC64;

Query Match 58.1%; Score 50; DB 14; Length 735;
 Best Local Similarity 46.7%; Pred. No. 34;
 Matches 7; Conservative 2; Mismatches 2; Indels 4; Gaps 1;

OY 2 LRMPW---WPMRK 12
 :|||
 Db 1 MAMPWRRRWRMR 15

RESULT 9
 ID Q9DT80 PRELIMINARY; PRT: 49 AA.
 AC Q9DT80;
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE ORF1 (FRAGMENT).
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TYM9;
 RX MEDLINE-20568739; PubMed-11118348;
 RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
 Sai T., Sugai Y.;
 RT "TT virus mRNAs detected in the bone marrow cells from an infected
 individual."
 RT Biochem. Biophys. Res. Commun. 279:700-707(2000).
 RL EMBL: AB050449; BAB19930.1;
 DR NON_TER 49
 FT SEQUENCE 49 AA; 7225 MW; 1DA6F8F1AB69AA43 CRC64;

Query Match 57.0%; Score 49; DB 14; Length 49;
 Best Local Similarity 36.8%; Pred. No. 4.3;
 Matches 7; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

OY 2 LRMPW-----WPRRK 12
 : | | | | |
 DB 1 MAWTMMORRRRRPWRRR 19

RESULT 10
 ID 019573 PRELIMINARY; PRT; 467 AA.
 AC 019573;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, last sequence update)
 DT 01-JUN-2000 (TREMblrel. 14, last annotation update)
 DE SIMILARITY TO 9 AMINO ACID REPEATS IN GALACTOSE SPECIFIC LECTINS.
 GN F18G3.2.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94150718; PubMed-7906398;
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,
 Crahan M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
 Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
 Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
 Parsons J., Percy C., McMurray A., Mortimore B., O'Callaghan M.,
 Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,
 Thierly-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
 Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 elegans."
 RT Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Favello T.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Waterston R.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U39855; AAA81062.1;
 KW Lectin.
 SQ SEQUENCE 467 AA; 53169 MW; 7D9BBAB61830431B CRC64;

Query Match 57.0%; Score 49; DB 5; Length 467;
 Best Local Similarity 83.3%; Pred. No. 31;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 WPMWPM 9
 : | | | |
 DB 201 WPMWPM 206

RESULT 11
 ID 09DT81 PRELIMINARY; PRT; 748 AA.
 AC 09DT81;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, last annotation update)
 DE ORF1
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TYM9;
 RX MEDLINE-20568739; PubMed-11118348;
 RA Okamoto H., Nishizawa T., Tawara A., Takahashi M., Kishimoto J.,
 Sai T., Sugai Y.;
 RT "TT virus mRNAs detected in the bone marrow cells from an infected
 individual."
 RT Biochem. Biophys. Res. Commun. 279:700-707(2000).
 RL EMBL: AB050449; BAB19928.1;
 DR EMBL: AB050449; BAB19928.1;
 SQ SEQUENCE 748 AA; 88352 MW; D65CCB2CAA5CE26F CRC64;

Query Match 57.0%; Score 49; DB 14; Length 748;
 Best Local Similarity 36.8%; Pred. No. 47;
 Matches 7; Conservative 2; Mismatches 2; Indels 8; Gaps 1;

OY 2 LRMPW-----WPRRK 12
 : | | | | |
 DB 1 MAWTMMORRRRRPWRRR 19

RESULT 12
 ID 09X8C2 PRELIMINARY; PRT; 114 AA.
 AC 09X8C2;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, last annotation update)
 DE HYPOTHETICAL 13.0 KDA PROTEIN.
 GN SCE36.09.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Oliver K., Harris D.;
 RC STRAIN-A3(2);
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RT Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE-97000351; PubMed-8843436;
 RA Redenbach M., Kieser H.M., Denaplatte D., Eichner A., Cullum J.,
 Kinsahl H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 RT Mol. Microbiol. 21:77-96(1996).

DR EMBL: AL049763; CAB42078.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 114 AA; 13031 MW; B2BE223FC4A0DBA9 CRC64;

Query Match 56.4%; Score 48.5; DB 2; Length 114;
 Best Local Similarity 80.0%; Pred. No. 10;
 Matches 8; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 OY 3 RW-PWMPWR 11
 11 111111
 DB 103 RWRPWRPWR 112

RESULT 13
 ID 007504 PRELIMINARY; PRT; 540 AA.
 AC 007504;
 DT 01-JUL-1997 (TREMblrel. 04, Created)
 DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
 DT 01-AUG-1998 (TREMblrel. 07, Last annotation update)
 DE GERMINATION PROTEIN.
 GN GERA.
 OS Bacillus megaterium.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OX NCBI_TaxID=1404;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 12872;
 RA Tani K., Nasu M., Ishitobi Y.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U61380; AAB61393.1;
 SQ SEQUENCE 540 AA; 59764 MW; 33A66ABFC13491FC CRC64;

Query Match 55.8%; Score 48; DB 2; Length 540;
 Best Local Similarity 58.3%; Pred. No. 47;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 ILRPMWPMWRK 12
 11111111
 DB 491 LRLPMWPMWR 502

RESULT 14
 ID 09SNN3 PRELIMINARY; PRT; 165 AA.
 AC 09SNN3;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)
 DE HYPOTHETICAL PROTEIN.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 OC Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae;
 OC Oryza
 OX NCBI_TaxID=4530;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC
 clone: P0493C11."
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AP000559; BAA84796.1;
 SQ SEQUENCE 165 AA; 18358 MW; 9EE82A556ED22DE9 CRC64;

Query Match 54.7%; Score 47; DB 10; Length 165;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 PMWPM 9
 11111
 DB 40 PMWPM 44

RESULT 15
 ID 09HXC9 PRELIMINARY; PRT; 276 AA.
 AC 09HXC9;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
 DE PROBABLE SHORT-CHAIN DEHYDROGENASE.
 GN PA3883.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=287;
 RN 111
 RP SEQUENCE FROM N.A.
 RC STRAIN-PA01;
 RA MEDLINE-20437337; PubMed-10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,
 Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL: AE004805; AAC07270.1;
 DR InterPro: IPR002198;
 DR InterPro: IPR002347;
 DR Pfam: PF00106; adh_short; 1.
 DR PRINTS: PR00081; GDRHDM.
 DR PRINTS: PR00080; SDRPFAMLY.
 DR PROSITE: PS00061; ADH_SHORT; UNKNOWN_1.
 SQ SEQUENCE 276 AA; 29427 MW; C47458ACCA15A3 CRC64;

Query Match 54.7%; Score 47; DB 2; Length 276;
 Best Local Similarity 70.0%; Pred. No. 35;
 Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 RWPMPWRK 12
 11111111
 DB 197 RWPMPWRK 206

Search completed: June 21, 2001, 08:37:44
 Job time: 246 sec

